

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 18, 2005, 13:24:07 ; Search time 15065 Seconds
(without alignments)

12792.525 Million cell updates/sec

Title: US-09-931-375A-1

Perfect score: 5063

Sequence: 1 gccatgagagccgagtgagc.....agggctggagagacttgta 5063

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues
Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

BEST: *
1: gb_est1: *
2: gb_est2: *
3: gb_hic: *
4: gb_est3: *
5: gb_est4: *
6: gb_est5: *
7: gb_est6: *
8: gb_gest1: *
9: gb_gest2: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	* Query Match	Length	DB	ID	Description
1	2059	40.7	2838	3	AK090382	AK090382 Mus muscu
2	1365.8	27.0	2752	3	AK081025	AK081025 Mus muscu
3	932	18.4	1045	5	BM920216	BM920216 AGENCOURT
4	919.6	18.2	1032	5	BM920216	BM920216 AGENCOURT
5	914.8	18.1	1031	5	BM920216	BM920216 AGENCOURT
6	877.8	17.3	1052	5	BM920216	BM920216 AGENCOURT
7	836.4	16.5	888	5	BM920216	BM920216 AGENCOURT
8	811	16.0	967	5	BM920216	BM920216 AGENCOURT
9	807	15.9	870	5	BM920216	BM920216 AGENCOURT
10	805	15.9	1001	5	BM920216	BM920216 AGENCOURT
11	802.2	15.8	986	4	BM06571	BM06571 AGENCOURT
12	785.2	15.5	932	4	BM06571	BM06571 AGENCOURT
13	784.4	15.5	902	5	BM06571	BM06571 AGENCOURT
14	783.4	15.5	855	5	BM06571	BM06571 AGENCOURT
15	781.6	15.4	5400	9	BM06571	BM06571 AGENCOURT
16	779	15.4	911	5	BM06571	BM06571 AGENCOURT
17	778.4	15.4	891	5	BM06571	BM06571 AGENCOURT
18	776.8	15.3	866	5	BM06571	BM06571 AGENCOURT
19	771.6	15.2	5400	9	BM06571	BM06571 AGENCOURT
20	766.4	15.1	880	7	BM06571	BM06571 AGENCOURT
21	762.8	15.1	902	5	BM06571	BM06571 AGENCOURT
22	758	15.0	886	6	BM06571	BM06571 AGENCOURT
23	732	14.5	847	5	BM06571	BM06571 AGENCOURT
24	723.6	14.3	5400	9	BM06571	BM06571 AGENCOURT

c 25	717.2	14.2	877	5	BM068854	BM068854 AGENCOURT
c 26	711	14.0	876	5	BM068854	BM068854 AGENCOURT
c 27	707.4	14.0	943	5	BM068854	BM068854 AGENCOURT
c 28	704.4	13.9	706	5	BM068854	BM068854 AGENCOURT
c 29	702.8	13.9	776	4	BM068854	BM068854 AGENCOURT
c 30	702.6	13.9	1026	2	BM068854	BM068854 AGENCOURT
c 31	699.4	13.8	834	1	BM068854	BM068854 AGENCOURT
c 32	696.6	13.8	995	4	BM068854	BM068854 AGENCOURT
c 33	688.2	13.6	956	5	BM068854	BM068854 AGENCOURT
c 34	687.6	13.6	712	5	BM068854	BM068854 AGENCOURT
c 35	684.4	13.5	821	1	BM068854	BM068854 AGENCOURT
c 36	677.6	13.4	860	4	BM068854	BM068854 AGENCOURT
c 37	673	13.3	1212	4	BM068854	BM068854 AGENCOURT
c 38	671.8	13.3	921	4	BM068854	BM068854 AGENCOURT
c 39	670.8	13.2	1076	2	BM068854	BM068854 AGENCOURT
c 40	661.8	13.1	699	7	BM068854	BM068854 AGENCOURT
c 41	652.6	12.9	854	5	BM068854	BM068854 AGENCOURT
c 42	645.2	12.7	919	6	BM068854	BM068854 AGENCOURT
c 43	636.8	12.6	850	5	BM068854	BM068854 AGENCOURT
c 44	636.2	12.6	655	6	BM068854	BM068854 AGENCOURT
c 45	635.6	12.6	938	2	BM068854	BM068854 AGENCOURT

ALIGNMENTS

RESULT 1	AK090382	2838 bp	mRNA	linear	HTC 03-APR-2004
LOCUS	AK090382				
DEFINITION	Mus musculus adult male intestinal mucosa cDNA, RIKEN full-length enriched library, clone:G630070K03 product:low density lipoprotein receptor-related protein 5, full insert sequence.				
ACCESSION	AK090382				
VERSION	AK090382.1	GI:26105912			
KEYWORDS	HTC; CAP trapper.				
ORGANISM	Mus musculus (house mouse)				
REFERENCE	1 Carrincci, P. and Hayashizaki, Y. High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999)				
JOURNAL	99279253				
MEDLINE	10349636				
REFERENCE	2 Carrincci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komo, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)				
AUTHORS	11042159				
TITLE	Shibata, K., Itoh, M., Atzawa, K., Nagaoka, S., Sasaki, N., Carrincci, P., Komo, H., Akiyama, J., Nishi, K., Kitsuana, T., Teshiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multichannel sequencer Genome Res. 10 (11), 1757-1771 (2000)				
JOURNAL	20530913				
MEDLINE	11076861				
REFERENCE	4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)				
JOURNAL	5 The FANTOM Consortium and the RIKEN Genome Exploration Research				
REFERENCE					
AUTHORS					

TITLE Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE (bases 1 to 2838)
AUTHORS Adachi,J., Aikawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hirooka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Kato,H., Kawai,U., Kojima,Y., Kondo,S., Konno,H., Konda,M.,
Koyama,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Nunazaki,R., Ohno,M., Ohsato,N.,
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shingawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku,Akashira,S.,
Takeda,Y., Tanaka,T., Tomaru,A., Toyota,T., Yasunishi,A.,
Muzumatsu,M. and Hayashizaki,Y.
TITLE Direct Submission
JOURNAL Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@sc.riken.jp,
URL: http://genome-gsc.riken.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
COMMENT cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in Riken.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome-gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/
location/Qualifiers
1..2838
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM:DB:G630070K03"
/db_xref="taxon:10090"
/clone="G630070K03"
/sex="male"
/tissue_type="intestinal mucosa"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
1..2644
/note="Low density lipoprotein receptor-related protein 5
[MGI:MGI:1278315, GB|NM_008513, evidence: BLASTN, 99%,
match=2144]
putative"
misc_feature
Query Match 40.7%; Score 2059; DB 3; Length 2838;
Best Local Similarity 85.4%; Pred. No. 0;
Matches 2424; Conservative 0; Mismatches 375; Indels 40; Gaps 10;
ORIGIN
2230 TGGATGGGCAAGACCTTACTAGGCGGACACTGGGACCAAGATGAGTGGCGGG 2289
|||||
Db 2 TGGATGGGCAAGACCTTACTAGGCGGACCAAGATGAGTGGCGGG 61
2290 CTGACGGGCAAGTCCGCGCAAGTCTCGTGTGAGGAGCTTGACCAACCGAGTGGT 2348
|||||
Db 62 CTGATGGGCAAGTCCGCGCAAGTCTCGTGTGAGGAGCTTGACCAACCGAGTGGT 121
2349 GGGCTTGATCCCAAGGCTACATCTGACCGAGTGGGGCGGCAAGCGAGAT 2408
|||||
Db 122 GGGCTTGATCTCTAAAGCTACATCTGACCTGAGTGGGGCGGCAAGGAT 181
2409 CGTGGGGCTTCAAGGCGGCAAGCTGATGACGCTGTGACCAAGT -GGGCGGG 2467
|||||
Db 182 TGTGGGGCTTCAAGGCGGCAAGCTGATGACGCTGTGACCAAGTGGGGCGG 241
2468 CCAAGCACTTCAAGCTGACCTGACGAGCGCTCTACTGACCGACTGACCA 2527

Db 242 CCAAGCACTTCAAGCTGACCTGACGAGCGCTCTACTGACCGACTGACCA 301
Qy 2528 ACATGATCGAGTGGTCCCAACATGCTGGGTGAGAGCGGGTGGTATGGCGAGATCTCC 2587
Db 302 ACATGATGAGTGGTCCCAACATGCTGGGTGAGAGCGGGTGGTATGGCGAGATCTCC 361
Qy 2588 GGCACCCGTTGGTGTGAGCGAGTACAGCGATTAATATCTACTGACAGACTGGAATCTGC 2647
Db 362 CTTACCCGTTGGTGTGAGCGAGTACAGCGATTAATATCTACTGACAGACTGGAATCTGC 421
Qy 2648 ACAGATGAGCGGGCGGCAAGATGACCGCGGGAACCGGACCTCTACCGGGCCAC 2707
Db 422 ATAGATGAGCGGGCGGCAAGATGACCGCGGGAACCGGACCTCTACCGGGGTCA 481
Qy 2708 TGAATGATGAGCAATCTGATGTTTCACTCTCCCGCGAGGATGGCTCAATGACT 2767
Db 482 TGAATGATGAGCAATCTGATGTTTCACTCTCCCGCGAGGATGGCTCAATGACT 541
Qy 2768 GTATGCAACAACGGGCAAGTGTGGGCACTGTGCTTGGCATCCCGGGCGGCAAGCT 2827
Db 542 GGTGCAACAAGTGTGGGCAAGTGTGGGCACTGTGCTTGGCATCCCGGGCGGCAAGCT 601
Qy 2828 GCGGCTGGGCTCAACTACAC -CGTGAACCGGCAAGCGGCAAGTGAAGCGGGCCAC 2886
Db 602 GTGGCTGGCTTCAACTACACCGGCTGAACCGGCAAGCGGCAAGTGAAGCGGGCCAC 653
Qy 2887 ACCTTCTGTGTTCAAGCAAGAACTGTCATGATGATGATGATGATGATGATGATGATGAT 2946
Db 654 -----AAATTCATGACCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 692
Qy 2947 AGCCGATTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3006
Db 693 AGCCGATTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 752
Qy 3007 CCACTGCAAGATTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3066
Db 753 CCGCTGCAAGATTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 812
Qy 3067 GACGGGACCAAGCCCTTGTGTTTGAACCTCTGAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAG 3126
Db 813 GACGGGACCAAGCCCTTGTGTTTGAACCTCTGAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAG 872
Qy 3127 CCGGACCAAGCCCTTGTGTTTGAACCTCTGAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAG 3186
Db 873 CCGGACCAAGCCCTTGTGTTTGAACCTCTGAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAG 932
Qy 3187 AATACATCAAGCTTCAAGGCTGAGCGGGAAGCAATGGGGGTGTTGCTGTTGGGGAG 3246
Db 933 AATACATCAAGCTTCAAGGCTGAGCGGGAAGCAATGGGGGTGTTGCTGTTGGGGAG 992
Qy 3247 CCGGCAAGCCCAAGGCTGAGCGGGAAGCAATGGGGGTGTTGCTGTTGGGGAG 3306
Db 993 CCGGCAAGCCCAAGGCTGAGCGGGAAGCAATGGGGGTGTTGCTGTTGGGGAG 1052
Qy 3307 ATGAGAGACCGGGAAGCAAGATGGAAGCGGCAAGCCCTTGAAGCGGCAAGCGGGAAGT 3366
Db 1053 ATGAGAGACCGGGAAGCAAGATGGAAGCGGCAAGCCCTTGAAGCGGCAAGCGGGAAGT 1112
Qy 3367 CTTTCAACACCGGCTCATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3426
Db 1113 CTTTCAACACCGGCTCATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1172
Qy 3427 CTGTTCTGGGTGAGCGGCAAGCTGAGCGGATGAGCGGCTGAGCGGCGGCAAG 3486
Db 1173 CTGTTCTGGGTGAGCGGCAAGCTGAGCGGATGAGCGGCTGAGCGGCGGCAAG 1230
Qy 3487 CGCTGACCTTGAAGAGCGCAACATGCTGAGCTCTGGGCTGACCATCTTGGCAAG 3546
Db 1231 CGCTGACCTTGAAGAGCGCAACATGCTGAGCTCTGGGCTGACCATCTTGGCAAG 1290
Qy 3547 CATCTACTGATGACCGGCAAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 3606

Db 1291 CACCTCTACTGATCGACCGCCAGCAGCAGATGATCGAGCGCGTGGAGAACCACTCGG 1350
Qy 3607 GACAAAGCGACTCGCATCCAGGCGCGTGTGCGCCACCTCCTCAGTGGCATCATGCACTGAG 3666
Db 1351 GACAAAGCGACTAGGGTTTCAAGGCGCGTGTGCGCCACCTCAGTGGCATCATGCGCT -GAG 1409
Qy 3667 GAAATCAGCTGAGGAGTTTTCAGCCCAACCATGTGCGCGTGTGCAATGGTGGTGTCTCC 3726
Db 1410 GAAATCAGCTGAGGAGTTTTCAGCCCAACCATGTGCGCGTGTGCAATGGTGGTGTCTCC 1469
Qy 3727 CACATCTGTATTTGCAAGGTTGATGGGACACACAGTGTCTATGCGCCAGTCCACCTCTG 3786
Db 1470 CACATCTGTATGCGCAAGGTTGATGGGACACACAGTGTCTATGCGCCAGTCCACCTCTG 1529
Qy 3787 CTCCTGAGAACCTCTGCTGACCTGTGAGAGCGCGCCACCTGCTCCCGGACAGTTTGCA 3846
Db 1530 CTCCTGAGAACCTCTGCTGACCTGTGAGAGCGCGCTCTACTGCTCCCTGATCAGTTTGA 1589
Qy 3847 TGTGCCACAGGGAGATCGACTGTATCCCGGGGCTGTGCGTGTGAGCGGCTTTTCCCGAG 3906
Db 1590 TGTACCACTGGTGAATCGACTGTATCCCGGAGCTGTGAGCGGCTTCCCTCTGAG 1649
Qy 3907 TGCATGACAGAGCAGAGAGGCTGCGCGGTGTGCTGCGCGCCAGTTCCTCCCTGC 3966
Db 1650 TGTGCTGACCAAGGTTGATGAAGAGGCTGCCAGTGTGCTCGCGCTCTCAGTTCCCTCTGC 1709
Qy 3967 GCGCGGGTGTGCTGAGCTGCGCTGCGCTGCGCTGCGAGCGGAGGAGTGTGAGGAG 4026
Db 1710 GCTGAGGCCAGTGTGAGAGCTGCGGTGTGAGCTGCGAGCGGAGGAGTGTGAGGAG 1769
Qy 4027 CGCTCAGAGCGGCGGACTGTGAGCGCCATCTGCTGCGCCCAACAGTTCGGGTGTGCGAGC 4086
Db 1770 CGCTCTGATGAAGCTAACTGGAGTGTGCTGCTGCTGCGCCCAACAGTTCGGGTGTGCGAGC 1829
Qy 4087 GCGAGTGTGCTCATCAACAGAGTGTGAGTCTTCCCGGAGTGTGAGCGGCTTTTCCCGAG 4146
Db 1830 GCGAGTGTGCTCATCAACAGAGTGTGAGTCTTCCCGGAGTGTGAGCGGCTTTTCCCGAG 1889
Qy 4147 GACGAGCTCATGTGTGAATCACCAAGCGCGCTCAGAGCAGCGCGCGCCACAGCAGT 4206
Db 1890 GATGAGCTCATGTGTGAATCAACAGCGCACCTCTGATGACATCCCGGAGCGAGCAGT 1949
Qy 4207 GCCATCGGCGCGCTCATTTGGCATCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4266
Db 1950 GCCATCGGCGCGCTCATTTGGTATCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2009
Qy 4267 GTGTGCGAGCGGTGTGCGAGCGCTATGCGGGGCGCAACAGCGGCTTTTCCCGAGCAG 4326
Db 2010 GTCTGCCAGCGTGTGATGCGAGCGCTACACAGGGGCGAGTGGGCGCTTTTCCCGAGCAG 2069
Qy 4327 TATGTCAGGGGACCCCGCAGCGTGGCCCTCAATTTATAGCCCGGGCGGTTCAGCAT 4386
Db 2070 TATGTTGTGAGGCGCTCATGTGCTCTCACTCTCACTCTCACTCTCACTCTCACTCTCA 2129
Qy 4387 GGC -CCCTTTCAGGCGATCGATGCGGAAAGTCCATGATGAGTCCGCTGAGCGCTGATGG 4445
Db 2130 GGCTCCCTTCCAGGATCCCGTGCAGCAGTCCGCTGATGAGTCCATGAGCGCTGGTGG 2189
Qy 4446 GGGCGGGGCGGGGTG -CCCTGTACGACCGGAAACCATGTCAGAGGGCGCTCTGTCAGCA 4504
Db 2190 GGGGCGGCGAGTGTGCGCCCTCTATGACCGGAATCACTGCTGAGCGCTCATCCAGCA 2249
Qy 4505 GCTGTTCAGCAGCAGGCGAGCTGTACCGCGCGATCTGAAACCGCGCGCTTCCCGG 4564
Db 2250 GCTGTTCAGCAGCAGGCGAGCTGTATATCGCGCGATCTGAAACCGCGCGCTTCCCGG 2309
Qy 4565 CCAGGACCTCTCTGTGACAACTGAGCATGTTCTACTTCTTCAAACTTCGCGCACTG 4624
Db 2310 CCAGGACCTCTCTGTGACAACTGAGCATGTTTATTTCTTTCAGGATCCCGGCGAG 2369
Qy 4625 CGAGCAGTACAGGCGCTTACATCTTCAGGAGTGGCGCGCGCGCGCGCGCTTTCAGCA 4684
Db 2370 CTAGAGCATACAGGCGCTTACGTCATTCGAGGTATGGCACCGCGCGCGCGCGCGCT 2429

Qy 4685 CCGAGCTGTGTGACAGCGACTACAGCGCTGGAAGCGCCAGCAAGTACTACCTGG 4744
Db 2430 CAGATGTGTGTGACAGTGTGACTACAGCAGTGTGCTGAGAGCAGCAAACTACTACCTGG 2489
Qy 4745 ATTGAACTCGGACTCAGACCCCTTATCCACCCCAACCCCAACCCCAACCCCAACCTGT 4804
Db 2490 ACTTGAATTCGACTCAGACCCCTTATCCACCCCAACCCCAACCCCAACCCCAACCTAT 2549
Qy 4805 CGCGGAGGAGCAGTGTGCGCGCTGCGCGCCAGCAGGAGGCTACTTCCATCTCTTCC 4864
Db 2550 CTGAGAGGAGCAGTGTGCGCGCTGCGCGCCAGCAGGAGGCTACTTCCATCTCTTCC 2609
Qy 4865 GCGCCCTCTGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTTCT 4924
Db 2610 GCGCCCTCTGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTG 2667
Qy 4925 CTGTGCGCTCTGAAATAGTGTGAAATAGTGTGAAATAGTGTGAAATAGTGTGAAAT 4984
Db 2668 CTGCTCTCTGAAATAGTGTGAAATAGTGTGAAATAGTGTGAAATAGTGTGAAATAGT 2726
Qy 4985 AAAATAAATAAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGA 5044
Db 2727 AAAATAAATAAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGA 2786
Qy 5045 GGGCTGGGAGAACTTTGTA 5063
Db 2787 GGGCTGGGAGAACTTTGTA 2805

RESULT 2

AK081025

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

AK081025 2752 bp mRNA linear HTC 03-APR-2004
Mus musculus 10 days neonate cerebellum cDNA, RIKEN full-length
enriched library, clone:B93069L06 product:low density lipoprotein
receptor-related protein 5, full insert sequence.

AK081025
HTC; CAP trapper.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636

2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159

3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multipipette sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

4
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.


```
QY 1318 CTCGACTGGTGGCCGCGAAACCTCTACTGGACCGACGCGGACCGGACCGCATCGAGTG 1377
Db 1388 GTGACTGGTGGCCGCGAAACCTCTACTGGACGAGATCAGGACCTGACAGATTGAGTG 1447
QY 1378 AGCGCTCTCAACGGGACCTCCGCGAAGATCTCTGTTGTCGAGGACCTGGAAGCGCCCGA 1437
Db 1448 ACTCGCTCTCAACGGGACCTCCGCGAAGATCTCTGTTATCTGAGGACCTGGAAGCGCCCGA 1507
QY 1438 GCCATCGCAC-TGCACCCCGGTGATGGGCTCATGTTACTGGACGACACTGGGAGAGAACCC 1496
Db 1508 GCCATTGGTTGACCTCTGATGGGCTCATGTTACTGGACGACTGGGAGAGAACCC 1567
QY 1497 TAAATCGAGTGTGCCAACTTGGATGGGACGAGCGGCTGTGTTGTTCAATGCTCCCT 1556
Db 1568 CAAATCGAATGGCCCAACCTAGATGGGAGAGATCGGCATGTCCTGTTGAACACTCCCT 1627
QY 1557 CGGTGGCCCAACGGCTGGCCCTGGACCTGAGAGGGGAACTCTACTGGGAGAGCG 1616
Db 1628 TGGTGGCCCAACGGACTGGGCTGGACCTGAGAGGGGAACTCTACTGGGAGAGCG 1687
QY 1617 CAAGACAGACAAGATCGAGGTGATCAATGTTGATGGGACGAGGCGGACCTCCCTGGA 1676
Db 1688 CAA-ACTGATAAATCGAGGTGATCAATGATGGGACGAGGCGGAAAGCCCTGCTTGA 1746
QY 1677 GGACAAGCTCCCGCACATTTTGGGTTTCAAGCTGCTGGGAGGACTTCTACTGAGCTGA 1736
Db 1747 GGACAAGCTCCCGCACATTTTGGGTTTCAAGCTGCTGGGAGGACTTCTACTGAGCGGA 1806
QY 1737 CTGGAGCGCGGACGAGTCGAGCGGCTGCAAGGTCAAGGCGGACGCGGACGTCATCAT 1796
Db 1807 CTGGAGAGAGCGAGTATTTGAAAGGGTCCCAAGGTCAAGGCGGACGCGGAGTGTATCAT 1866
QY 1797 TGACCAAGTCCCGGACCTGAGGGGCTCAAGCTGTGAATGTGCCAAGGTGCTCGG 1853
Db 1867 TGATCAACTCCCGGACCTGAGGGGACTCAAGGCTGTGAATGTGCCAAGGTGCTCGG 1923

RESULT 3
BM920216
LOCUS BM920216 1045 bp mRNA linear EST 12-MAR-2002
DEFINITION AGENCOURT_6706353 NIH_MGC_120 Homo sapiens cDNA clone IMAGE:5750120
5', mRNA sequence.
ACCESSION BM920216
VERSION BM920216.1 GI:19370595
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12780 row: j column: 09
High quality sequence start: 29
High quality sequence stop: 736.
Location/Qualifiers
1. 1045
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5750120"
/lab_host="DH10B"
/clone_lib="NIH_MGC_120"
```

```
ORIGIN
Query Match 18.4%; Score 932; DB 5; Length 1045;
Best Local Similarity 96.6%; Pred. No. 1.1e-192;
Matches 972; Conservative 0; Mismatches 32; Indels 2; Gaps 2;
QY 3202 CACAGGCTGAGCGGGGAGCCATCGGGGTGTGCTGCTGGGACCGGACGAGCCGAG 3261
Db 41 CCCAGGCTGAGCGGGGAGCCATCGGGGTGTGCTGCTGGGACCGGACGAGCCGAG 100
QY 3262 GCCATCGTGTCAACGCGGAGCGAGGTTACTGTACTTTCACCAACATCGAGGACCGGCA 3321
Db 101 GCCATCGTGTCAACGCGGAGCGAGGTTACTGTACTTTCACCAACATCGAGGACCGGCA 160
QY 3322 GCCAAGATCGAAGCGGACGCTTGGACCGGACCGGAGGCTCTTTCACCAACCGGC 3381
Db 161 GCCAAGATCGAAGCGGACGCTTGGACCGGAGCGGAGGCTCTTTCACCAACCGGC 220
QY 3382 CTGATCGCCCTGTGGCCCTGGTGTGGACAAACACATCGGCGAAGCTGTTCTGGGTGAC 3441
Db 221 CTGATCGCCCTGTGGCCCTGGTGTGGACAAACACATCGGCGAAGCTGTTCTGGGTGAC 280
QY 3442 GCGGACCTGAAGCGCATTTGAGAGCTGTGACCTGTGAGGGGCAACCGGCTGACCTGAG 3501
Db 281 GCGGACCTGAAGCGCATTTGAGAGCTGTGACCTGTGAGGGGCAACCGGCTGACCTGAG 340
QY 3502 GACGCCAACAATCGTGCAGCCTCTGGGCTTGACCAATCTTGGCAAGCATCTTCTACTGGATC 3561
Db 341 GACGCCAACAATCGTGCAGCCTCTGGGCTTGACCAATCTTCTACTGGATC 400
QY 3562 GACGCCAGCAGCAGATGATCGAGCTGTGGAGAAAGACACCGGGGACAGGAGCTGCG 3621
Db 401 GACGCCAGCAGCAGATGATCGAGCTGTGGAGAAAGACACCGGGGACAGGAGCTGCG 460
QY 3622 ATCCAGGCGCGTGTGCGCCACCTCACTGGCATTCATGAGTGGAGGAGTCAAGCTGAG 3681
Db 461 ATCCAGGCGCGTGTGCGCCACCTCACTGGCATTCATGAGTGGAGGAGTCAAGCTGAG 520
QY 3682 GAGTTCTCAGCCCAACCATGTCGCCGTGACAAATGGTGGCTGCTCCCAATCTGTATTGCC 3741
Db 521 GAGTTCTCAGCCCAACCATGTCGCCGTGACAAATGGTGGCTGCTCCCAATCTGTATTGCC 580
QY 3742 AAGGCTGATGGGACACACACGCTGCTCATGCCAGTCCACCTGCTGCTGAGGAGCTG 3801
Db 581 AAGGCTGATGGGACACACACGCTGCTCATGCCAGTCCACCTGCTGCTGAGGAGCTG 640
QY 3802 CTGACCTGTGGAGAGCGCCGACCTGCTCCCGGACAGTGTGATGTCACACAGGCGAG 3861
Db 641 CTGACCTGTGGAGAGCGCCGACCTGCTCCCGGACAGTGTGATGTCACACAGGCGAG 700
QY 3862 ATCGACTGTATTCCCGGGGCTGCGCTGTGACGCGCTTCCGAGTGGATGACAGAGC 3921
Db 701 ATCGACTGTATTCCCGGGGCTGCGCTGTGACGCGCTTCCGAGTGGATGACAGAGC 760
QY 3922 GACGAGAGGCTGCTCCCGGTGCTGCTCCCGGCGGAGTCCCTGCGCGGGGTCAGTGT 3981
Db 761 GACGAGAGGCTGCTCCCGGTGCTGCTCCCGGCGGAGTCCCTGCGCGGGGTCAGTGT 820
QY 3982 GTGACCTGTGCGCTGCGCTGCGAGCGGAGGAGCAGACTGTGAGGACCGCTCAGACAGGCG 4041
Db 821 GTGACCTGTGCGCTGCGCTGCGAGCGGAGGAGCAGACTGTGAGGACCGCTCAGACAGGCG 880
QY 4042 GACTGTGACGCGCATCTGCTGCCCAACCATGTTCCGGTGTGGAGGCGGCGAGTGTCTCTC 4101
```

/note="Organ: pooled pancreas and spleen; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of spleen and pancreas from 28 yo male. Library is oligo-dr primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-2.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 025. Note: this is a NIH_MGC Library."

```

Db      881  GACTGTGACGNCATCTGCTG-CAAAACANTTCCGGTGTGCCAGCGGCATGTGTCTCTC 939
Qy      4102 ATCAAAACAGCAGTCGACTCCTTCCCGACTGTATCGAGCGGTCCGACGAGCTCATGTGT 4161
Db      940  ATCAAAC-GCAGTGGCGGTCTCTTCCCGACTGTATCGAGCGGTCCGACGAACTCAGTGGG 998
Qy      4162 GAAATCACCAAGCGCCCTCAGACGACAGCCCGGCCACAGCAGTG 4207
Db      999  GAAATCACAGCGCCCTCAGAACGAGCCCGGCCCAACAGTG 1044

RESULT 4
LOCUS    BX439521
DEFINITION BX439521 1032 bp mRNA linear EST 04-MAY-2004
5-PRIME, mRNA sequence.
ACCESSION BX439521
VERSION    BX439521.2 GI:47020909
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1032)
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 15, 2003 this sequence version replaced gi:30787803.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 6245.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna/s=CS0DE011BG01QPI&c=6245.f.
Location/Qualifiers
1..1032
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE011IM02"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

FEATURES
source
Query Match 18.2%; Score 91.9.6; DB 5; Length 1032;
Best Local Similarity 99.4%; Pred. No. 5.6e-190;
Matches 930; Conservative 3; Mismatches 2; Indels 1; Gaps 1;

Qy 1323 CTGGGTGGCCGAAACCTCTACTGACGACACGCGGACCGCATCGAGGTGACGCG 1382
Db 1 CTGGGTGGCCGAAACCTCTACTGACGACACGCGGACCGCATCGAGGTGACGCG 60
Qy 1383 CCTCAACGGACCTCCCGCAAGATCTTGGTTCGAGACCTCGAGACCGCCGAGCCAT 1442
Db 61 CCTCAACGGACCTCCCGCAAGATCTTGGTTCGAGACCTCGAGACCGCCGAGCCAT 120
Qy 1443 CGCACTGCACCCCTGTATGGCCCTCATGTACTGACAGACTGGGAGAACCTTAAAT 1502
Db 121 CGCACTGCACCCCTGTATGGCCCTCATGTACTGACAGACTGGGAGAACCTTAAAT 180
Qy 1503 CGAGTGTGCCAACTTGGATGGCAGGACCGCGTGTGCTGTGCTCAATGCTCCCTCGGGTG 1562

```

```

Db      181  CGAGTGTGCCAACTTGGATGGCAGGAGCGCGTGTGCTGCTCAATGCTCCTCCCTCGGGTG 240
Qy      1563 GCCCAACGGCCCTGGCCCTTGGACCTTGCAGGAGGGGAGAGCTTACTTGGGGAGAGCGCAAGAC 1622
Db      241  GCCCAACGGCCCTGGCCCTTGGACCTTGCAGGAGGGGAGAGCTTACTTGGGGAGAGCGCAAGAC 300
Qy      1623 AGACAAGATCGAGTGCATGATGTTGATGGACGAGAGGGGGAGCCCTCCTTGGAGGACAA 1682
Db      301  AGACAAGATCGAGTGCATGATGTTGATGGACGAGAGGGGGAGCCCTCCTTGGAGGACAA 360
Qy      1683 GCTCCCGCACATTTTCGGGTTTCAGCTGCTGGGGGACTTTCATCTACTGAGTGCATGCGCA 1742
Db      361  GCTCCCGCACATTTTCGGGTTTCAGCTGCTGGGGGACTTTCATCTACTGAGTGCATGCGCA 420
Qy      1743 GCGCGGACGATCGAGCGGGTGCACAAGGTCAAGGCCAGCGGGAGCGTGCATCATTTGACCA 1802
Db      421  GCGCGGACGATCGAGCGGGTGCACAAGGTCAAGGCCAGCGGGAGCGTGCATCATTTGACCA 480
Qy      1803 GCTGCCCGACCTGATGGGGTTCAAAGCTGTCAAGTGTGGCCCAAGGTTCGTCGGAACCAACCC 1862
Db      481  GCTGCCCGACCTGATGGGGTTCAAAGCTGTCAAGTGTGGCCCAAGGTTCGTCGGAACCAACCC 540
Qy      1863 GTGTGCGGACAGGAACCGGGGTGCAGCCACTGTGCTTCTTCACACCCCAACGCAACCCG 1922
Db      541  GTGTGCGGACAGGAACCGGGGTGCAGCCACTGTGCTTCTTCACACCCCAACGCAACCCG 600
Qy      1923 GTGTGGCTGCCCATCGGCTTGGAGTGTCTGAGTGACATGAAAGACCTGTCATCGTGCCTGA 1982
Db      601  GTGTGGCTGCCCATCGGCTTGGAGTGTCTGAGTGACATGAAAGACCTGTCATCGTGCCTGA 660
Qy      1983 GGCCTTCTTGTGCTTCCACAGCAGAGCGCCATCCACAGGATCTCCCTCGAGAGCAATAA 2042
Db      661  GGCCTTCTTGTGCTTCCACAGCAGAGCGCCATCCACAGGATCTCCCTCGAGAGCAATAA 720
Qy      2043 CAACGAGCTGGCCATCCCGTTCACGGGCGTCAAGAGGCGCTCAGCCCTGGACTTTTGATGT 2102
Db      721  CAACGAGCTGGCCATCCCGTTCACGGGCGTCAAGAGGCGCTCAGCCCTGGACTTTTGATGT 780
Qy      2103 GTCCAAACACACATCTACTTGGACAGAGCTGAGCTGAAAGACCATCAGCCGCGCTTCAT 2162
Db      781  GTCCAAACACACATCTACTTGGACAGAGCTGAGCTGAAAGACCATCAGCCGCGCTTCAT 840
Qy      2163 GAACGGAGCTCGGTGGAGCAGCTGGTGGAGTGGCTTGGCTTACCTACCCCGAGGGCATGCG 2222
Db      841  GAACGGAGCTCGGTGGAGCAGCTGGTGGAGTGGCTTGGCTTACCTACCCCGAGGGCATGCG 900
Qy      2223 CGTTGACTGGATGGGCAAGAACCTCTACTGGGCGCA 2258
Db      901  CGTTGACTGGATGGGCAAA-AACTCTACTGGGCGCRM 935

RESULT 5
BX402084/c
LOCUS    BX402084
DEFINITION BX402084 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
cDNA clone CS0DC011YB06 3-PRIME, mRNA sequence.
ACCESSION BX402084
VERSION    BX402084.2 GI:46874938
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1031)
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 13, 2003 this sequence version replaced gi:30618689.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime

```


QY 2557 CAGGACGGGTGCTGATTTGGGACGATCTCCGACACCGTTGCGTCTGACGAGTACAGC 2616
 Db |||||
 421 CAGGACGGGTGCTGATTTGGGACGATCTCCGACACCGTTGCGTCTGACGAGTACAGC 480
 QY 2617 GATTATATCTACTGACAGACTGGAATCTGCACAGACTGAGCGGCGGACAGACTAGC 2676
 Db |||||
 481 GATTATATCTACTGACAGACTGGAATCTGCACAGACTGAGCGGCGGCGGACAGACTAGC 540
 QY 2677 GCGCGGAACCGCACCTCATCCAGGGCCACCTGGACTTCTGATGAGACATCTGCTGTTTC 2736
 Db |||||
 541 GCGCGGAACCGCACCTCATCCAGGGCCACCTGGACTTCTGATGAGACATCTGCTGTTTC 600
 QY 2737 CACTCTCCCGCCAGGATGGCTCAATGACTGTATGCAACAACAGGGCAGTGTGGGCGAG 2796
 Db |||||
 601 CACTCTCCCGCCAGGATGGCTCAATGACTGTATGCAACAACAGGGCAGTGTGGGCGAG 660
 QY 2797 CTGTGCTTGGCATCCCGGGGCGGACCGCTGCGGCTGCGCTCACACTACACCTGGAC 2856
 Db |||||
 661 CTGTGCTTGGCATCCCGGGGCGGACCGCTGCGGCTGCGCTCACACTACACCTGGAC 720
 QY 2857 CCCAGCAGCCGCACTGCGAGCCCGCCACACCTTCTTGTGTTTTCAGCCAGAAATCTGCC 2916
 Db |||||
 721 CCCAGCAGCCGCACTGCGAGCCCGCCACACCTTCTTGTGTTTTCAGCCAGAAATCTGCC 780
 QY 2917 ATCAGTCGATGATCCCGGACGACGACGACGCCC-GGATCTCATCTCTGCGCCCTGCATGG 2975
 Db |||||
 781 ATCAGTCGATGATCCCGGACGACGACGACGCCC-GGATCTCATCTCTGCGCCCTGCATGG 840
 QY 2976 ACTGAGGACGTCGAAGCCTGAGTATGACC 3007
 Db |||||
 841 GACTGAGAACGTCGAAGCCTGAGTATGACC 872

RESULT 8
 BX428434
 LOCUS BX428434 Homo sapiens PLACENTA Homo sapiens linear EST 05-MAY-2004
 DEFINITION 5-PRIME, mRNA sequence.
 ACCESSION BX428434
 VERSION BX428434.1 GI:31017882
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 967)
 Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
 was not normalized. Library was constructed by Life Technologies, a
 division of Invitrogen.
 This sequence belongs to sequence cluster 6245.f
 For more information about this cluster, see
 http://www.genoscope.cns.fr/cdna?e=CS0A006ZH04_A0512_1&c=6245.f.
 Location/Qualifiers
 1. 967
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DE011YM02"
 /tissue_type="PLACENTA"
 /clone_lib="Homo sapiens PLACENTA"
 /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
 with a NotI-oligo(dT) primer. Five prime end enriched,
 double-strand cDNA was digested with Not I and cloned into
 the Not I and EcoRV sites of the pCMVSPORT 6 vector."

ORIGIN
 Query Match 16.0%; Score 811; DB 5; Length 967;
 Best Local Similarity 90.3%; Pred. No. 3.1e-166;
 Matches 859; Conservative 0; Mismatches 92; Indels 0; Gaps 0;
 Library was not normalized."

QY 1874 GGAACGGGGGTGACGACCACTGCTCTTCTTCAACACCCACGCAACCCGGTGTGCTGCGC 1933
 Db |||||
 6 GGTACGGGGGTGACGACCACTGCTCTTCTTCAACACCCACGCAACCCGGTGTGCTGCGC 65
 QY 1934 CCATCGGCTCGAGCTGCTGAGTGACATGAAGACCTGCTGCTGAGGCTTCTTTGG 1993
 Db |||||
 66 CCATCGGCTCGAGCTGCTGAGTGACATGAAGACCTGCTGCTGAGGCTTCTTTGG 125
 QY 1994 TCTTCAACGACGAGCGGCCATCCACAGGATCTCCCTCGAGACCAATAAACGACGTGG 2053
 Db |||||
 126 TCTTCAACGACGAGCGGCCATCCACAGGATCTCCCTCGAGACCAATAAACGACGTGG 185
 QY 2054 CCATCCCGCTCAACGGGGTCAAGGAGGCTCAGGCCCTGGACTTTTGTATGTGTCCAAACAC 2113
 Db |||||
 186 CCATCCCGCTCAACGGGGTCAAGGAGGCTCAGGCCCTGGACTTTTGTATGTGTCCAAACAC 245
 QY 2114 ACATCTACTTGAACAGAGCTCAGCCCTGAAGACCATCAGCCGCGCTTTCATGAACGGGAGCT 2173
 Db |||||
 246 ACATCTACTTGAACAGAGCTCAGCCCTGAAGACCATCAGCCGCGCTTTCATGAACGGGAGCT 305
 QY 2174 CGGTGGAGCAGCTGGTGGAGTTTGGCCTTGAATACCCCGAGGCGATGCGCTTGAATGA 2233
 Db |||||
 306 CGGTGGAGCAGCTGGTGGAGTTTGGCCTTGAATACCCCGAGGCGATGCGCTTGAATGA 365
 QY 2234 TGGGCAAGAACTCTTACTTGGGCCGACACTGGGACCAATAGAAATCGAAGTGGCGCGCTGG 2293
 Db |||||
 366 TGGGCAAGAACTCTTACTTGGGCCGACACTGGGACCAATAGAAATCGAAGTGGCGCGCTGG 425
 QY 2294 ACGGCGAGTTCCGGCAAGTCTCTGTGTGAGGAGCTTTGGACAAACCCGAGGTGCTGTGGCCC 2353
 Db |||||
 426 ACGGCGAGTTCCGGCAAGTCTCTGTGTGAGGAGCTTTGGACAAACCCGAGGTGCTGTGGCCC 485
 QY 2354 TGGATCCCAACGAGGCTACATCTACTTGGACCGAGTGGGGCGGACGAGGATCGTGC 2413
 Db |||||
 486 TGGATCCCAACGAGGCTACATCTACTTGGACCGAGTGGGGCGGACGAGGATCGTGC 545
 QY 2414 GGGGCTTTCATGAGCGGACCAACTGCACTGCTGTGGACAAAGTGGCGCGGCGCAACG 2473
 Db |||||
 546 GGGGCTTTCATGAGCGGACCAACTGCACTGCTGTGGACAAAGTGGCGCGGCGCAACG 605
 QY 2474 ACCTCACCATTTGACTACGCTGACGAGCGCTTCTACTGACCGGACCTGGACCAACATGA 2533
 Db |||||
 606 ACCTCACCATTTGACTACGCTGACGAGCGCTTCTACTGACCGGACCTGGACCAACATGA 665
 QY 2534 TCGAGTCTGCAACATGCTGGTTCAGAGCGGGTCTGATTTGCCGACGATCTCCCGGAC 2593
 Db |||||
 666 TCGAGTCTGCAACATGCTGNGTTCAGAGCGGGTCTGATTTGCAAGGATCTCCGCGAC 725
 QY 2594 CGTTGCGTTCGACCGCAGTACAGCGATATATCTACTGACAGACTGGAATCTGCACAGCA 2653
 Db |||||
 726 CGTTGCGTTCGACCGCAGTACAGCGATATATTTCTACTGACAGACTGGAATCTGCACAGCA 785
 QY 2654 TTGAGCGGGCGGACCAAGACTAGCGCGCGGAAACCGCACCTCATCCAGGCGCACCTGGACT 2713
 Db |||||
 786 TTGAGCGGGCGGACCAAGACTAGCGCGCGGAAACCGCACCTCATCCAGGNNCCCTGNACT 845
 QY 2714 TCGTGTGAGCATCTCTGTGTTTCCATCTCTCCGCGAGGATGGGCTCAATGACTGTATGC 2773
 Db |||||
 846 NCGTGTGAGCATCTCTGTGTTTCCCTCTCTNNCCGAGNAGNCCTCATGACTGTGTCACAC 905
 QY 2774 ACAACAAAGGCGCAGTGTGGGCGAGCTGTGCCCTTTGCCATCCCGGGCGGCCACC 2824
 Db |||||
 906 TCGGCACTGTGGCAGTTGGGCTTTCCTTTTCGCGGACCCGCTGCGNTTAGC 956

RESULT 9

BQ687102/c
LOCUS BQ687102 870 bp mRNA linear EST 15-JUL-2002
DEFINITION AGENCOURT_8046500 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6209180
5', mRNA sequence.
ACCESSION BQ687102
VERSION BQ687102.1 GI:21812406
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1. (bases 1 to 870)
TITLE NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2368 row: a column: 21
High quality sequence stop: 680.
FEATURES
source
1..870
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6209180"
/tissue_type="ductal carcinoma, cell line"
/lab_host="PH10B (phage-resistant)"
/clone_lib="NIH MGC 110"
/note="Organ: Pancreas; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; CDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGACAGG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
ORIGIN
Query Match 15.9%; Score 807; DB 5; Length 870;
Best Local Similarity 98.7%; Pred. No. 2.3e-165;
Matches 824; Conservative 0; Mismatches 10; Indels 1; Gaps 1;
3396 GGCCCTGGTGGACACACACT-GGGCAAGCTGTTCTGGTGGACGGCGACCTGAAGC 3454
835 GCCCGTGGTGGACACACACTGGGGCAAGCTGTTCTGGGGACGGCGGACCTGAAGC 776
3455 GCATTGAGACTGTGACCTGTACAGGGGCCAACCGCTGACCTTGGAGGACGCCAATCG 3514
775 GCATTGAGACTGTGACCTGTACAGGGGCCAACCGCTTGAACCTGGAGGACGCCAATCG 716
3515 TGACGCTCTGGGCTGACCATCTTGGGAGCATCTTACTGATCGACCGCCAGCAGC 3574
715 TGACGCTCTGGGCTGACCATCTTGGGAGCATCTTACTGATCGACCGCCAGCAGC 656
3575 AGATGATCGAGCGTGTGGAGAGACACACCGGGGACAGCGGACTCGCATCCAGGGCCGCTG 3634
655 AGATGATCGAGCGTGTGGAGAGACACACCGGGGACAGCGGACTCGCATCCAGGGCCGCTG 596
3635 TCGCCCACTCTACCTGGCATCCATGAGTGGAGAGTGCACCTGGAGAGTTCAGGCC 3694
595 TCGCCCACTCTACCTGGCATCCATGAGTGGAGAGTGCACCTGGAGAGTTCAGGCC 536
3695 ACCCATGTCGCCGAGCAATGGTGGCTGCTCCACATCTCTATTGCAAGGGGTGATGGGA 3754
535 ACCCATGTCGCCGAGCAATGGTGGCTGCTCCACATCTCTATTGCAAGGGGTGATGGGA 476

3755 CACACCGGTGCTCATGCCACCTCCAGCTCCTCTCCAGAACCTGCTGACCTGTGGAG 3814
475 CACACCGGTGCTCATGCCACCTCCAGCTCCTCTCCAGAACCTGCTGACCTGTGGAG 416
3815 AGCCGCCACCTGCTCCCGGACACAGTTTGCATGTGCCACAGGGGAGATCGACTATCC 3874
415 AGCCGCCACCTGCTCCCGGACACAGTTTGCATGTGCCACAGGGGAGATCGACTATCC 356
3875 CCGGGGCTGGCGCTGTGACCGCTTCCGAGTGCATGACGAGCGAGGAGGGCT 3934
355 CCGGGGCTGGCGCTGTGACCGCTTCCGAGTGCATGACGAGCGAGGAGGGCT 296
3935 GCCCGGTGCTCCCGCGCCAGTTCCTCCGCGCGGGGTGAGTGTGTGGACCTGCGCC 3994
295 GCCCGGTGCTCCCGCGCCAGTTCCTCCGCGCGGGGTGAGTGTGTGGACCTGCGCC 236
3995 TGCGCTGCGACGCGGAGGAGACTGTGACGACCGCTCAGACGAGCGGAGCTGTGAGCCCA 4054
235 TGCGCTGCGACGCGGAGGAGACTGTGACGACCGCTCAGACGAGCGGAGCTGTGAGCCCA 176
4055 TCTGCTGCGACACAGTTCCTCCGAGTGCATGCGGCGGCGGAGTGTCTCATCAACAGCAGT 4114
175 TCTGCTGCGACACAGTTCCTCCGAGTGCATGCGGCGGCGGAGTGTCTCATCAACAGCAGT 116
4115 GCGACTCTCTCCCGGAGTGTGACGACCGCTCAGACGAGCTCATGTGTGAAATCACCAGC 4174
115 GCGACTCTCTCCCGGAGTGTGACGACCGCTCAGACGAGCTCATGTGTGAAATCACCAGC 56
4175 GCGCTCTCAGACGAGCGCGGCCACAGCGAGTGCATCGGCGGCGGCTCATGTGGCAT 4229
55 GCGCTCTCAGACGAGCGCGGCCACAGCGAGTGCATCGGCGGCGGCTCATGTGGCAT 1
RESULT 10
BX402085 1001 bp mRNA linear EST 29-APR-2004
LOCUS BX402085 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
DEFINITION cDNA clone CS0DC011YB06 5-PRIME, mRNA sequence.
ACCESSION BX402085
VERSION BX402085.2 GI:46876774
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1. (bases 1 to 1001)
AUTHORS Li W.B., Gruber C., Jesses, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 13, 2003 this sequence version replaced gi:30622712.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Crémieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
6245.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?c=CS1AC003ZF04QPI&c=6245.f.
FEATURES
source
1..1001
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DC011YB06"
/tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V

ORIGIN sites of the pCMVSPORT 6 vector. Library was normalized."

Query Match 15.9%; Score 805; DB 5; Length 1001;
Best Local Similarity 96.3%; Pred. No. 6.5e-165;
Matches 873; Conservative 5; Mismatches 21; Indels 8; Gaps 5;

QY 2978 TGAGGAACGTCAGAACCCATGACATATGACCCCACTGGACAAGTTCTACTTACTGGGTGGATG 3037
DB 4 TGAGGAACGTCAGAACCCATGACATATGACCCCACTGGACAAGTTCTACTTACTGGGTGGATG 63

QY 3038 GCGCCGACAACTAAGCGAGCGCAAGGAGCGGACCGACCGCCCTTTGTTTGACCTCTC 3097
DB 64 GCGCCGACAACTAAGCGAGCGCAAGGAGCGGACCGACCGCCCTTTGTTTGACCTCTC 123

QY 3098 TGAGCAAGGCGCAAAACCCAGACGCGCCCGCCAGCTCAGCATCGACATCTACAGCC 3157
DB 124 TGAGCAAGGCGCAAAACCCAGACGCGCCCGCCAGCTCAGCATCGACATCTACAGCC 183

QY 3158 GGACACTGTTCTGGACGTCGAGGCGCACCAATACCATCAAGTCCACAGGCTGAGCGGG 3217
DB 184 GGACACTGTTCTGGACGTCGAGGCGCACCAATACCATCAAGTCCACAGGCTGAGCGGG 243

QY 3218 AAGCCATGGGGTGTGCTGCTGGGGACCGCGACAAGCCGAGGCGCATCGTGTGCAAG 3277
DB 244 AAGCCATGGGGTGTGCTGCTGGGGACCGCGACAAGCCGAGGCGCATCGTGTGCAAG 303

QY 3278 CGGAGCGAGGTGACCTGTACTTCCACATCGAGCGCGGCGAGCCCAAGATCGAACCG 3337
DB 304 CGGA--SAGGGTACCTGTACTTCCACATCGAGCGCGGCGAGCCCAAGATCGAACCG 361

QY 3338 CAGCCCTGGAGCGGACCGAGCGCGAGGTCCTCTTCCACACCGGCTCATCGCCCTGTGG 3397
DB 362 CA-CCCTGGACCGGACCGAGCGGAGGTCCTCTTCCACACCGGCTCATCGCCCTGTGG 420

QY 3398 CCTGTGTGGTGAACAACAACACTGGGCAAGCTGTTCTGGGTGAGCGCGACTGGAAGCGCA 3457
DB 421 CCTGTGTGGTGAACAACAACACTGGGCAAGCTGTTCTGGGTGAGCGCGACTGGAAGCGCA 480

QY 3458 TTGAGAGCTGTGACCTGTGCGGGCGCAACCGCTGACCTGGGAGCGCCCAACATCGTGC 3517
DB 481 TTGAGAGCTGTGACCTGTGCGGGCGCAACCGCTGACCTGGGAGCGCCCAACATCGTGC 540

QY 3518 AGCCTCTGGGCTGACCATCTTGGCAAGCATCTCTACTGATCGACCGCAGCAGCAGA 3577
DB 541 AGCCTCTGGGCTGACCATCTTGGCAAGCATCTCTACTGATCGACCGCAGCAGCAGA 600

QY 3578 TGATCGAGCGTGTGGAAGAACCACCGGGGACAAGCGGACTCGCATCCAGGGCGGTGTGC 3637
DB 601 TGATCGAGCGTGTGGAAGAACCACCGGGGACAAGCGGACTCGCATCCAGGGCGGTGTGC 660

QY 3638 CCCACCTCACTGGCATCCATGAGTGGAGGAGTCAAGCTGAGGAGGTTCTCAGGCCACC 3697
DB 661 CCCACCTCACTGGCATCCATGAGTGGAGGAGTCAAGCTGAGGAGGTTCTCAGGCCACC 720

QY 3698 CATGTGCCGTGACAAATGTTGCTCTCCACATCTGTATTGCAAGGCTGATGSGACAC 3757
DB 721 CATGTGCCGTGACAAATGTTGCTCTCCACATCTGTATTGCAAGGCTGATGSGACA- 779

QY 3758 CACGGTGTCTATGCCAGTCCACTGCTCTCTGAGAACTGCTGACCTGTGAGAGC 3817
DB 780 CACGGTGTCTATGCCAGTCCACTGCTCTCTGAGAACTGCTGACCTGTGAGAG- 838

QY 3818 GCGCCACCTGTCTCCCGGACCGATTGATGTCACAGGGAGATCGACTGTATCCCG 3877
DB 839 GCGCCACCTGTCTCCCGGACCGATTGATGTCACAGGGAGATCGACTGTATCCCG 895

QY 3878 GGGCCTG 3884
DB 896 GCTGCTG 902

RESULT 11

BM561582 986 bp mRNA linear EST 20-FEB-2002
LOCUS AGENCOURT 6567315 NIH_MGC_88 Homo sapiens cDNA clone IMAGE:5739763
DEFINITION 5', mRNA sequence.
ACCESSION BM561582
VERSION BM561582.1 GI:18060953
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 986)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Straubeberg, Ph.D.
Email: csapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/ILNL at:
http://image.llnl.gov
Plate: L1AM12753 row: j column: 20
High quality sequence stop: 680.

FEATURES
Location/Qualifiers
1..986
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5739763"
/tissue_type="duodenal adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_88"
/note="Organ: Small intestine; Vector: pCMV-SPORT6;
Site 1: NotI; Site 2: SalI; Cloned unidirectionally;
oligo-dT primed. Average insert size 1.767 kb. Library
enriched for full-length clones and constructed by Life
Technologies. Note: this is a NIH_MGC Library."

ORIGIN
Query Match 15.8%; Score 802.2; DB 4; Length 986;
Best Local Similarity 98.6%; Pred. No. 2.7e-164;
Matches 831; Conservative 0; Mismatches 8; Indels 4; Gaps 2;

QY 1778 CCAGCCGGAGCGTCATATTGACAGTGCCCGACCTGATGGGCTCAAAGCTGTGAATG 1837
DB 1 CCAGCCGGAGCGTCATATTGACAGTGCCCGACCTGATGGGCTCAAAGCTGTGAATG 60

QY 1838 TGGCCAAAGTGTGTCGGACCAACCCGTGTGCGGACAGGAACGGGGGTGCAGCCACCTGT 1897
DB 61 TGGCCAAAGTGTGTCGGACCAACCCGTGTGCGGACAGGAACGGGGGTGCAGCCACCTGT 120

QY 1898 GCTTCTTTCACACCCACGCAACCCGGTGTGCTGCCCATCGGCTGGAGCTGTGAGTG 1957
DB 121 GCTTCTTTCACACCCACGCAACCCGGTGTGCTGCCCATCGGCTGGAGCTGTGAGTG 180

QY 1958 ACATGAAGACCTGTCATGCTGAGGCTTTCTTGGTCTTCCACGAGAGCGCCCATCC 2017
DB 181 ACATGAAGACCTGTCATGCTGAGGCTTTCTTGGTCTTCCACGAGAGCGCCCATCC 240

QY 2018 ACAGGATCTCCCTCGAGACCAATAACACGAGCTGGCCATCCCGCTCAGCGGCTCAAGG 2077
DB 241 ACAGGATCTCCCTCGAGACCAATAACACGAGCTGGCCATCCCGCTCAGCGGCTCAAGG 300

QY 2078 AGGCTCTAGCCCTGGAATTTGATGTGTCTCAACAACACCATCTACTGGAACAGCTGACGC 2137
DB 301 AGGCTCTAGCCCTGGAATTTGATGTGTCTCAACAACACCATCTACTGGAACAGCTGACGC 360

QY 2138 TGAAGACCATCAGCCGCGCTTTCATGAACGGGAGCTCGGTGGAGCACTGTGTGAGTTTG 2197
DB 361 TGAAGACCATCAGCCGCGCTTTCATGAACGGGAGCTCGGTGGAGCACTGTGTGAGTTTG 420

QY 2198 GCCTTGAATACCCGAGGCGATGCGCGTTGACTGATGGCAAGAACCTCTACTGGCGG 2257
Db 421 GCCTTGAATACCCGAGGCGATGCGCGTTGACTGATGGCAAGAACCTCTACTGGCGG 480
QY 2258 ACACCTGGGACCAACAGAAATCGAAGTGGCGCGCTGGAGCGGCGAGTTCGCGCAAGTCCCTCG 2317
Db 481 ACACCTGGGACCAACAGAAATCGAAGTGGCGCGCTGGAGCGGCGAGTTCGCGCAAGTCCCTCG 540
QY 2318 TGTGGAGGAGCTTGACAAACCGAGGTCGTGGCCCTGGATCCCAAGGCGCTACATCT 2377
Db 541 TGTGGAGGAGCTTGACAAACCGAGGTCGTGGCCCTGGATCCCAAGGCGCTACATCT 600
QY 2378 ACTGGACCGAGTGGGCGCGCAAGCCGAGGATCGTGGCGGCTTCATGACGCGGACCAACT 2437
Db 601 ACTGGACCGAGTGGGCGCGCAAGCCGAGGATCGTGGCGGCTTCATGACGCGGACCAACT 660
QY 2438 GCATGACGCTGCTGGTGAACGAGTGGCGCGGCGCAACGACCTCACCAATGACTACGCTGACC 2497
Db 661 GCATGACGCTGCTGGTGAACGAGTGGCGCGGCGCAACGACCTCACCAATGACTACGCTGACC 720
QY 2498 AGCGCTCTACTGGACCGACCTGGACACCAACATGATCGAGTGGTCCAAACATGCTGGG 2555
Db 721 AGCGCTCTACTGGACCGACCTGGACACCAACATGATCGAGTGGTCCAAACATGCTGGG 780
QY 2556 TCAGGAGCGGTCGATGTCGCGAGTCTCCGCA--CCGTTGCGTCTGACGCGATAC 2613
Db 781 CAGGAGCGGTCGATGTCGCGAGTCTCCGCAACCGTTCGGGCGCTGACGCGATAC 840
QY 2614 AGC 2616
|||
Db 841 AGC 843
|||

RESULT 12

BM006571/c
LOCUS
DEFINITION 932 bp mRNA linear EST 30-OCT-2001
603615546F1 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:5421467 5',
mRNA sequence.

ACCESSION BM006571

VERSION BM006571.1

KEYWORDS GI:16520925

SOURCE EST.

ORGANISM Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 932)

NIH-MGC <http://mgs.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLCM1877 row: 1 column: 12

High quality sequence stop: 757.

FEATURES

source

1..932

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5421467"

/tissue_type="ductal carcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_110"

/note="Organ: pancreas; Vector: pOTB7; Site: 1: XhoI;

Site 2: EcoRI; cDNA made by oligo-dT priming.

Directionally cloned into EcoRI/XhoI sites using the

following 5' adaptor: GGACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN

Query Match Similarity 15.5%; Score 785.2; DB 4; Length 932;
Best Local Similarity 96.5%; Pred. No. 1.4e-160;
Matches 867; Conservative 0; Mismatches 23; Indels 8; Gaps 6;
QY 3337 GCAGCCCTGGAGCGCACCGAGCGAGGTCTCTTACCACCGGCTCATCGCCCTGTG 3396
Db 895 GCCAGCTGGAGCGCAAGCGCGGCTCGTGGTCAACACCGG-CTCAACCGCCCTGTG 837
QY 3397 GCCCTGGTGGTGAACAACACTGGGCAAGCT- GTTCTGGGTGACGCGGACCTGAAAGCG 3455
Db 836 G-CCTGGTGGTGGCAACACACTGGGCAAGCTGTCTTCTGGGGGACGCGGACCTGAAGCG 778
QY 3456 CATTGAGAGCTGTGACCTGTTCAGGGGCCAACCGGCTGACCTGGAGGACG-CCACATCG 3514
Db 777 CATTGAGA--CAGGACCTGTTCAGGGGCCAACCGGCTGACCTGGAGGACGCGCCAACTCG 720
QY 3515 TGCA--GCCTCTGGGCTTGACCATCTTGGCAAGCATCTCTACTGGATCGACCGCCAGCA 3572
Db 719 TGAAGCTCTGGGCTTGACCATCTTGGCAAGCATCTCTACTGGATCGACCGCCAGCA 660
QY 3573 GCATGATCGAGCGTGTGGAGAGACCAACCGGGACAAAGCGGACTCGCATCCAGGGCGG 3632
Db 659 GCATGATCGAGCGTGTGGAGAGACCAACCGGGACAAAGCGGACTCGCATCCAGGGCGG 600
QY 3633 TGTGGCCCACTCACTGGCATCCATGAGTGGAGGAGTCAAGCTGGAGGATGTTCTCAGC 3692
Db 599 TTTGGCCCACTCACTGGCATCCATGAGTGGAGGAGTCAAGCTGGAGGATGTTCTCAGC 540
QY 3693 CCACCATATGCCCGCTGACATGCTGTCTCCCAATCTGTATTGCAAGGGTGATGG 3752
Db 539 CCACCATATGCCCGCTGACATGCTGTCTCCCAATCTGTATTGCAAGGGTGATGG 480
QY 3753 GACACACGCGTGTATGCCCGAGTCCACCTCGTGTCTCTGAGAACTCTGTGACTGTGG 3812
Db 479 GACACACGCGTGTATGCCCGAGTCCACCTCGTGTCTCTGAGAACTCTGTGACTGTGG 420
QY 3813 AGAGCGCCCACTGCTCCCGGACCAAGTTGATGTCACACAGGAGGATGCTGAT 3872
Db 419 AGAGCGCCCACTGCTCCCGGACCAAGTTGATGTCACACAGGAGGATGCTGAT 360
QY 3873 CCCCAGGCGCTGGCGCTGTGACGCGCTTTCCGAGTGGGATCACAGAGCGAGGAGGG 3932
Db 359 CCCCAGGCGCTGGCGCTGTGAGCGCTTCCGAGTGGGATCACAGAGCGAGGAGGG 300
QY 3933 CTGCCCCGTGTCTCCCGCGCCCAAGTTCCTCTGCGCGGGGTCAGTGTGTGGACCTGCG 3992
Db 299 CTGCCCCGTGTCTCCCGCGCCCAAGTTCCTCTGCGCGGGGTCAGTGTGTGGACCTGCG 240
QY 3993 CCTGCGCTGGAGCGGCGGAGGAGCTGTCAAGGACCGCTCAAGAGGCGGAGTGTGACGC 4052
Db 239 CCTGCGCTGGAGCGGCGGAGGAGCTGTCAAGGACCGCTCAAGAGGCGGAGTGTGACGC 180
QY 4053 CATCTGCTGCCCAACCAAGTTCCGGTGTGGAGCGGCGAGTGTCTCTCATCAACAGCA 4112
Db 179 CATCTGCTGCCCAACCAAGTTCCGGTGTGGAGCGGCGAGTGTCTCTCATCAACAGCA 120
QY 4113 GTGCGACTCTCTTCCCGACTGTATCGACGGTCCGACGAGCTCATGTGTGAATCACCAG 4172
Db 119 GTGCGACTCTCTTCCCGACTGTATCGACGGTCCGACGAGCTCATGTGTGAATCACCAG 60
QY 4173 GCCGCCCTCAGACGACGCGCGCCACAGAGTCCATCGGGCCCTCATTTGGCATC 4230
Db 59 GCCGCCCTCAGACGACGCGCGCCACAGAGTCCATCGGGCCCTCATTTGGCATC 2

RESULT 13

```
BU195422
LOCUS       BU195422       902 bp      mRNA      linear      EST 04-SEP-2002
DEFINITION AGENCOURT 7593499 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6171960
5' mRNA sequence.
ACCESSION  BU195422
VERSION     BU195422.1 GI:22709406
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 902)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13541 row: c column: 01
            High quality sequence stop: 638.

FEATURES             source
     source
     1..902
     /organism="Homo sapiens"
     /mol_type="mRNA"
     /db_xref="taxon:9606"
     /clone="IMAGE:6171960"
     /tissue_type="retinoblastoma"
     /lab_host="DH10B (phage-resistant)"
     /clone_lib="NIH_MGC_67"
     /note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;
            Site 2: SalI; Cloned unidirectionally. Primer: Oligo
            Average insert size 1.75 kb. Library constructed by Life
            Technologies."

ORIGIN
Query Match      15.5%; Score 784.4; DB 5; Length 902;
Best Local Similarity 95.6%; Pred. No. 2e-160;
Matches 862; Conservative 0; Mismatches 31; Indels 9; Gaps 5;

QY 3520 CCTCTGGGCGCTGACCATCTTGGCAGCATCTCTACTGATCGACCGCAGCAGATG 3579
Db 1 CCACGCGTCCGACCATCTTGGCAGCATCTCTACTGATCGACCGCAGCAGATG 60

QY 3580 ATCGAGCGGTGGAGAAGACCAACCGGGACAGCGGACTCGCATCCAGGCGGTGTGCC 3639
Db 61 ATCGAGCGGTGGAGAAGACCAACCGGGACAGCGGACTCGCATCCAGGCGGTGTGCC 120

QY 3640 CACCTCACTGGCATCCATGCACTGGAGGAAGTCAGCTGGAGGAGTTCTCAGCCACCCA 3699
Db 121 CACCTCACTGGCATCCATGCACTGGAGGAAGTCAGCTGGAGGAGTTCTCAGCCACCCA 180

QY 3700 TGTGCCCCGTGACATGTTGGCTGCTCCACATCTGTATTGTCAGAGGTGATGGACACCA 3759
Db 181 TGTGCCCCGTGACATGTTGGCTGCTCCACATCTGTATTGTCAGAGGTGATGGACACCA 240

QY 3760 CGGTGCTCATGCCAGTCCACTCGTGTCTCTGCAGAACCTGTCTGACCTGTGGAGAGCG 3819
Db 241 CGGTGCTCATGCCAGTCCACTCGTGTCTCTGCAGAACCTGTCTGACCTGTGGAGAGCG 300

QY 3820 CCCACTGCTCCCCGACCAAGTTTCATGTGTGCCACAGGGGAGATGACTGTATCCCCGGG 3879
Db 301 CCCACTGCTCCCCGACCAAGTTTCATGTGTGCCACAGGGGAGATGACTGTATCCCCGGG 360

QY 3880 GCCTGGCGCTGTGACGGCTTTCCCGAGTGCATGACCAAGGACGAGGAGGCTGCCCC 3939
Db 361 GCCTGGCGCTGTGACGGCTTTCCCGAGTGCATGACCAAGGACGAGGAGGCTGCCCC 420
```

```
3940 GTGTGCTCCGCGCGCCAGTTCCTCTGCGCGGGGTGAGTGTGTGAGACTGCGCCTGGCC 3999
Db 421 GTGTGCTCCGCGCGCCAGTTCCTCTGCGCGGGGTGAGTGTGTGAGACTGCGCCTGGCC 480

QY 4000 TCGACGCGCGAGGAGAGTGTCAAGACCGCTCAGACGAGGCGGACTGTGAGCCCATTCGC 4059
Db 481 TCGACGCGCGAGGAGAGTGTCAAGACCGCTCAGACGAGGCGGACTGTGAGCCCATTCGC 540

QY 4060 CTGCCCCAACCAAGTTCGGGTGCGAGCGGCGGAGTGTCTCTCAACACAGCAGTGGCAG 4119
Db 541 CTGCCCCAACCAAGTTCGGGTGCGAGCGGCGGAGTGTCTCTCAACACAGCAGTGGCAG 600

QY 4120 TCCTTCCCGGACTGTATCGAGCGGCTCCGACGAGCTCATGTGTGAATCACCAGCGGCC 4179
Db 601 TCCTTCCCGGACTGTATCGAGCGGCTCCGACGAGCTCATGTGTGAATCACCAGCGGCC 660

QY 4180 TCAGACGACAGCCCGGCCACAGCAGTGCCTATCGGGCCCGTCAATTGGCATCATCTCTCT 4239
Db 661 TCAGACGACAGCCCGGCCACAGCAGTGCCTATCGGGCCCGTCAATTGGCATCATCTCTCT 720

QY 4240 CTCCTCGTCATGGGTGGTCTATTTTGTGTGCCAGCG-CGTGGTGTGCCAGCGTATGC 4298
Db 721 CTCCTCGTCATGGGTGGTCTATTTTGTGGTGTGCCAGCGCGGGGTGTGCCAACGATGC 780

QY 4299 GGGGG--CCAAACGGGCGCTTCCCGCACAGTATGTACAGCGGACCCCGCAC---GTGCC 4353
Db 781 GGGGGCCAAACGGGCGCTTCCCGCACAGTATGTACAGCGGACCCCGCACGCGGCGGCC 840

QY 4354 CTCAAATTCATA-GCCCCGGGCGGTTCACAGCAT--GGCCCCCTTCACAGGATCGCATGC 4410
Db 841 CTCAAATTCATAACCCCGCGCGGTTCACAGCATTTCCAGCGGGACCCCGGACGATGCATGG 900

QY 4411 GG 4412
Db 901 GG 902

RESULT 14
BO228206
LOCUS      BO228206      855 bp      mRNA      linear      EST 02-MAY-2002
DEFINITION AGENCOURT 7593499 NIH_MGC_70 Homo sapiens cDNA clone IMAGE:6020654
5' mRNA sequence.
ACCESSION  BO228206
VERSION     BO228206.1 GI:20409606
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 855)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13225 row: b column: 15
            High quality sequence stop: 608.

FEATURES             source
     source
     1..855
     /organism="Homo sapiens"
     /mol_type="mRNA"
     /db_xref="taxon:9606"
     /clone="IMAGE:6020654"
     /tissue_type="epithelioid carcinoma"
     /lab_host="DH10B (phage-resistant)"
     /clone_lib="NIH_MGC_70"
```

/note="Organ: pancreas; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.1 kb. Library constructed by Life
Technologies."

ORIGIN

```

Query Match      15.5%; Score 783.4; DB 5; Length 855;
Best Local Similarity 97.9%; Pred. No. 3.3e-160;
Matches 836; Conservative 0; Mismatches 12; Indels 6; Gaps 4;

QY 3070 GGGACCCAGCCCTTGTGTTGACCTCTCTGAGCAGGCGCAAAACCCAGACAGCAGGCC 3129
DB 1 GGGACCCAGCCCTTGTGTTGACCTCTCTGAGCAGGCGCAAAACCCAGACAGGAGGCC 60

QY 3130 CACGACCTCAGCATCGACATCTACAGCCGAGCAGCTGTTCTGGAGTGTGGAGCCACCAAT 3189
DB 61 CACGACCTCAGCATCGACATCTACAGCCGAGCAGCTGTTCTGGAGTGTGGAGCCACCAAT 120

QY 3190 ACCATCAAGCTCCAGAGCTGAGCGGGGAGGAGCATGGGGGTGGTCTCGTGGGACCGC 3249
DB 121 ACCATCAAGCTCCAGAGCTGAGCGGGGAGGAGCATGGGGGTGGTCTCGTGGGACCGC 180

QY 3250 GACAGCCAGGGCCATCGTGTCAACCGGAGCGAGGTACTCTACTTCAACCAACATG 3309
DB 181 GACAGCCAGGGCCATCGTGTCAACCGGAGCGAGGTACTCTACTTCAACCAACATG 240

QY 3310 CAGGACCGGGCAGCCAAAGATCGAACCGCGAGCCCTTGGACCGCAGCGAGGTCTC 3369
DB 241 CAGGACCGGGCAGCCAAAGATCGAACCGCGAGCCCTTGGACCGCAGCGAGGTCTC 300

QY 3370 TTCACACCGGCTCATCGGCCCTGTGGCCCTGGTGTGGACACACATGGGCAAGCTG 3429
DB 301 TTCACACCGGCTCATCGGCCCTGTGGCCCTGGTGTGGACACACATGGGCAAGCTG 360

QY 3430 TTCTGGTGGACCGGACTGAAGCGCATTTGAGAGCTGTGACCTCTCAGGGGCCAACCGC 3489
DB 361 TTCTGGTGGACCGGACTGAAGCGCATTTGAGAGCTGTGACCTCTCAGGGGCCAACCGC 420

QY 3490 CTGACCTTGGAGGACCGCCAAATCGTGGAGCTCTTGGGCTGACCATCTTGGCAAGCAT 3549
DB 421 CTGACCTTGGAGGACCGCCAAATCGTGGAGCTCTTGGGCTGACCATCTTGGCAAGCAT 480

QY 3550 CTCTACTGATCGACCGCCAGCAGCAGATGATCGAGCGTGTGGAGAGACACACCGGGAC 3609
DB 481 CTCTACTGATCGACCGCCAGCAGCAGATGATCGAGCGTGTGGAGAGACACACCGGGAC 540

QY 3610 AAGCGGACTCGCATCCAGGGCGTGTGCGCCACCTCACTGGCATCCATGCAATGAGGAA 3669
DB 541 AAGCGGACTCGCATCCAGGGCGTGTGCGCCACCTCACTGGCATCCATGCAATGAGGAA 600

QY 3670 GTCAGCTGGAGGAGTTCTCAGCCACCATGTGCCCGTGACAAATGGTGGCTGCCAC 3729
DB 601 GTCAGCTGGAGGAGTTCTCAGCCACCATGTGCCCGTGACAAATGGTGGCTGCCAC 660

QY 3730 ATCTGATTTGCCAAGGTTGATGGACACCAACCGTGTCTATGCCAGTCCACCTCGTGCTC 3789
DB 661 ATCTGATTTGCCAAGGTTGATGGACACCAACCGTGTCTATGCCAGTCCACCTCGTGCTC 720

QY 3790 CTGCGAAGCTCTGTGACCTGTGGAGAGCGCCCACT-GCTCCCGGACCAATTTGCAATG 3848
DB 721 CTGCGAAGCTCTGTGACCTGTGGAGAGCGCCCACT-GCTCCCGGACCAATTTGCAATG 780

QY 3849 TGCC--ACAGGGAGATCGACTGTATCCCC--GGGGCTGGCGCTGTGACGCTTT-CCC 3903
DB 781 TGCC--ACAGGGAGATCGACTGTATCCCC--GGGGCTGGCGCTGTGACGCTTT-CCC 840

QY 3904 GAGTGGCATGACCA 3917
DB 841 GAGTGGCATGACCA 854

```

RESULT 15
AY407796

```

LOCUS      AY407796      5400 bp      DNA      linear      GSS 12-DEC-2003
DEFINITION Homo sapiens HCM3010 gene, VIRUAL TRANSCRIPT, partial sequence,
              genomic survey sequence.
ACCESSION   AY407796
VERSION     AY407796.1 GI:39763767
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE   1 (bases 1 to 5400)
AUTHORS     Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
              Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
              Ferrera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
              Adams,M.D. and Cargill,M.
TITLE       Inferring nonneutral evolution from human-chimp-mouse orthologous
              gene trios
JOURNAL     Science 302 (5652), 1960-1963 (2003)
PUBMED     14671302
REFERENCE   2 (bases 1 to 5400)
AUTHORS     Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
              Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
              Ferrera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
              Adams,M.D. and Cargill,M.
TITLE       Direct Submission
JOURNAL     Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
              Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
              them based on alignment.
FEATURES    Location/Qualifiers
             source          1..5400
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9606"
                        <1..5400
                        /locus_tag="HCM3010"
             gene           15.4%; Score 781.6; DB 9; Length 5400;
             origin          Best Local Similarity 54.1%; Pred. No. 1.2e-159;
                        Matches 1957; Conservative 0; Mismatches 1594; Indels 69; Gaps 15;

QY 147 GCCGCTCTCTGCTATTTCGCAACCGCCGGGAGCTACGGCTGGTGACCGCGCGGAGTCAA 206
DB 999 GCCTGTGCTGCTGTTCGCCAATCGCATCGGAGGTGCTGCC-----ACACCG 1052

QY 207 GCTGGAGTCCACCATCGTGTGTCAGCGGCTGGAGGATGCGGCCGAGTTCAGTTCAGTT 266
DB 1053 CTCTGAGTACACACTGCTGCTTAAACAACCTGGAGATGCCAATTGCCCTTGATTTCCACCA 1112

QY 267 TTCCAAGGAGCGCTGTACTGGACAGACGTGAGCGAGGAGGCCATCAAGCAGACCTACCT 326
DB 1113 CCGCGGAGGTTGTTCTTCTGTGTGATGTACCTCGACCGGATCTCCGTGCCAACCCT 1172

QY 327 GAACCCAGACGGGGCGCGCGTGCAGAAACGTGGTCACTCCGGCTGTGGTCTCTCCGACGG 386
DB 1173 CAAC-----GGCAGCAACGTGGAGGAGTTGTCTACTTGGCTGGAGAGCCCGAGGGG 1226

QY 387 CTTGCTGCTGCACTGGGTGGGCAAGAGCTGTACTGGACGAGCTCAGAGACCAACCGCAT 446
DB 1227 CTTGCTGCTGATTTGGGTGCCAATGCAAACTCTACTGGACCGACTCAGGCACCTCGAGAT 1286

QY 447 CGAGTGGCCCAACCTCAATGGCACATCCCGAAGGTGCTCTTTGGCAGGAGCTTTGACCA 506
DB 1287 TGAGTGGCCCAATCTGGACGGGGGCCACCGAGGTGTGCTGTGGCAGAGACCTGGAGAA 1346

QY 507 GCCTAGGGCCATCGCCTTTGGACCCCGCTCA CGGGTACATGTACTGGACAGACTGGGGTGA 566
DB 1347 GCGCCGGGCCATTCCTTTGCATCCCATGGAGGATACCAATTATTTCGACAGACTGGGGCAA 1406

QY 567 GACGCGCGGATTCGACGGGCGAGGATGGATGGGACGACCCCGAAGATCATTTGTGGACTC 626
DB 1407 CACCCCGGATTTGAGGCGCTTCAGCATGATGATGCTCTGACGCGCGCATCATTCGGATAC 1466

```


Db 3618 GTGGGCTCGAAGATGGCGGTGCTCCACCTCTGCTTGCTGGCCCTTCTGGCTTCTC 3677
Qy 2826 CTGGGCTGCGCTCACAATAACCTGGAGCCAGCAGCGCAACTGACGCCGCCAC 2885
Db 3678 CTGTGCTGCGCCCACTGAGCATCCAGCTGAAGGAGATGGGAAGACCTGTGATCCTCTCC 3737
Qy 2886 C---ACCTTCTTGTGTTTCAAGCAGAAATCTGCCATCAGTCGGATGATCCCGAGACCA 2942
Db 3738 TGAGACCTACCTGCTCTTCTCCAGCCGTGGGTCCATCCGGGTATCTCACTGGACACCAG 3797
Qy 2943 GCAGAGCCCGGATCTCATCTGCCCTCATGCTGAGGAGCTCAAGCCATCGACTA 3002
Db 3798 TGACACACCGATGTGATGCTCCCTGTTCTGAGCTCAACAATGTCATCTCCCTGGACTA 3857
Qy 3003 TGACCCACTGGACAAGTTTCACTACTGCGGTGGATGGCGCCAGAACATCAAGCCGAGCAA 3062
Db 3858 TGACAGCGTGGATGGAAAGGTCTATTACACAGATGTTCTCTGGATGTTA--TCAGGCGA 3915
Qy 3063 GGACGACGGGACCCAGCCCTTGTGTTTGAACCTCTTGAGCCAGGCAAAACCCAGACAG 3122
Db 3916 GCAGACCTGAACGGCAGCAACATGGAGACAGTGATCGGGCGAGGGCTGAAGACCA----- 3970
Qy 3123 GCAGCCCAACGACCTCAGCATCGACATCTACAGCGGACACTGTTCTGGACGTGGAGGC 3182
Db 3971 -----CTGACGGGCTGGCAGTGAGTGGGTGGCCAGGAACTGTACTGACAGACAGG 4025
Qy 3183 CACCAATACCATCAACGTCCACAGCTGAGCGGGAAAGCCATGGGGGTGGTGTGCGTGG 3242
Db 4026 TCGAAATACCATGAGGCGTCCAGGCTGGATGGTTCTCGCGCAAAAGTACTGATCAACAA 4085
Qy 3243 GGACCGCACAGCCAGGCCATCGTCTCAACGCGGAGCGGGTACTGTACTTCTAC 3302
Db 4086 TAGCCTGATGAGCCCGGCGCATGCTGTTTCCCGAGGAAGGGGTACTCTTCTGGAC 4145
Qy 3303 CAACATGACGAGCCGGCAGCAAGATCGAACGCGCAGCCTGGACGCGCACCGAGCGGA 3362
Db 4146 AGAC---TGGGGCCACATGTCAGATCGAACGGSCAACTTGGATGTTCTGAGCGGAA 4202
Qy 3363 GGTCTCTTCAACCGCGCTCATCCGCCCTGTGGCCCTGTGGTGGGACAAACACTGGG 3422
Db 4203 GGTCTCTATCAACACAGACCTGGGTGGCCCAATGGCCCTTACCCCTGGACTATGATACCCG 4262
Qy 3423 CAAGCTGTTCTGGGTGGACGGACCTGAGCGCATTTGAGGCTGTGACCTGTACAGGGC 3482
Db 4263 CAGGATCTACTGGGTGGATGGCATCTGGACCGGATCGAGAGTGTCTGACCTCAATGGG-- 4320
Qy 3483 CAACCGCTTGACCTGGAGGAGCGCAACATCGTGAGCCTCTGGGCCCTGACCATCTTGG 3542
Db 4321 -AACTGGCGCAGGTCTTGGTGGCCATGTGTCCACCCCTTGGCCCTCACACAGCAAGA 4379
Qy 3543 CAAGCATCTTACTGGATCGCCGCGCAGCAGAGATGATCGAGCGTGTGGAGAGACCAAC 3602
Db 4380 CAGGTGGATCTACTGGACAGACTGGCAGACCAAGTCAATCCAGCGTGTGACAAATATCTC 4439
Qy 3603 CGGGACAGCGGACTCGCATCCAGGGCGGTGTCGCCCACTCACTGGCATCCATGCAGT 3662
Db 4440 AGGCCGGAACAAGGAGACAGTGTCTGGCAATGTGAAGGACTCATGGATATC-----AT 4493
Qy 3663 GGAGGAAGTCAAGCTTGGAGGAGTCTTCAAGCCACCCATGTGCCCTGTGACAAATGGTGGCTG 3722
Db 4494 CGTGGTTTCCCTCAGCGCGCAGACAGGGACCAATGCTGTGTGTGAACAATGGTGGCTG 4553
Qy 3723 CTCCACATCTGTATTGCA 3742
Db 4554 CACCCACCTCTGCTTTGCCA 4573

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 18, 2005, 12:03:25 ; Search time 2493 Seconds
(without alignments)
12022.323 Million cell updates/sec

Title: US-09-931-375A-1

Perfect score: 5063

Sequence: 1 gccatggagcccgagtgagc.....aggctgggagaactttgta 5063

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N Geneseq_16Dec04:*

1: Geneseqn1980s:*

2: Geneseqn1990s:*

3: Geneseqn2000s:*

4: Geneseqn2001as:*

5: Geneseqn2001bs:*

6: Geneseqn2002as:*

7: Geneseqn2002bs:*

8: Geneseqn2003as:*

9: Geneseqn2003bs:*

10: Geneseqn2003cs:*

11: Geneseqn2003ds:*

12: Geneseqn2004as:*

13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5063	100.0	5063	6 AAD34322	Aad34322 Human bon
2	5061.4	100.0	5063	6 AAD34405	Aad34405 Human BSM
3	5061.4	100.0	5063	6 AAD34401	Aad34401 Human BSM
4	5061.4	100.0	5063	6 AAD34404	Aad34404 Human BSM
5	5058.2	99.9	5120	5 ABA82617	Abas2617 Human Zma
6	5058.2	99.9	5120	8 ACC45358	Acc45358 Human wil
7	5058.2	99.9	5120	10 ADB99038	Adb99038 Human LRP
8	5058.2	99.9	5120	10 ADE82639	Ade82639 Human Zma
9	5058.2	99.9	5120	13 ADRI6919	Adri6919 Human hig
10	5058.2	99.9	5278	8 ACC46079	Acc46079 Human ZMA
11	5058.2	99.9	5278	10 ADB98797	Adb98797 Human ZMA
12	5056.8	99.9	5099	13 ABD32937	Abd32937 Human can
13	5056.6	99.9	5120	6 ABK22777	Abk22777 Human hig
14	5056.6	99.9	5120	8 ACC45359	Acc45359 Human LRP
15	5056.6	99.9	5120	10 ADB99039	Adb99039 LRP5 vari
16	5056.6	99.9	5120	10 ADE82640	Ade82640 Human HBM
17	5056.6	99.9	5120	13 ADRI6920	Adri6920 Human hig
18	5056.6	99.9	5278	8 ACC46078	Acc46078 Mouse HBM
19	5056.6	99.9	5278	10 ADB98796	Adb98796 Mouse HBM
20	5055.2	99.8	5099	13 ADQ86437	Adq86437 Human tum

21	5055.2	99.8	5099	13 ADQ83350	Adq83350 Human tum
22	5055	99.8	5120	5 ABA82618	Abas2618 Human HBM
23	5052	99.8	5100	10 ADK66938	Adk66938 Gene #28
24	5052	99.8	5100	13 ADR73481	Adr73481 Human low
25	5050	99.7	5066	6 AAD34403	Aad34403 Human BSM
26	5036.6	99.5	5112	13 ADRI47571	Adri47571 Human hig
27	5032.2	99.4	5115	13 ADRI47570	Adri47570 Human hig
28	5028	99.3	5098	2 AAV70396	Aav70396 LRP5 isof
29	5010.8	99.0	5135	11 ACN44787	Acn44787 Human mRN
30	5010.8	99.0	5135	13 ABD32938	Abd32938 Human can
31	4999	98.7	5059	12 ADQ20523	Adq20523 Human sof
32	4952.6	97.8	5599	12 ADQ24509	Adq24509 Human sof
33	4902	96.8	5022	2 AAV85549	Aav85549 LRP5 isof
34	4894.4	96.7	5263	2 AAV70400	Aav70400 LRP5 isof
35	4890.6	96.6	5125	2 AAV86019	Aav86019 Lrp5 isof
36	4890	96.6	5165	2 AAV85551	Aav85551 LRP5 isof
37	4889.6	96.6	5162	2 AAV85550	Aav85550 LRP5 isof
38	4840	95.6	4848	12 ADI27199	Adi27199 Human LRP
39	4813	95.1	4843	2 AAV70395	Aav70395 LRP5 CDNA
40	4723.6	93.3	4915	2 AAV70398	Aav70398 LRP5 isof
41	4584.2	90.5	4736	6 ABK22776	Abk22776 Human Zma
42	4322.6	85.4	4351	2 AAV70397	Aav70397 LRP5 isof
43	3931	77.6	5119	6 ABA94317	Abas4317 LRP5 poly
44	3907	77.2	5117	2 AAV86018	Aav86018 Mouse Lrp
45	3900.2	77.0	5018	13 ABD32935	Abd32935 Mouse can

ALIGNMENTS

RESULT 1

AAD34322

ID AAD34322 standard; cDNA; 5063 BP.

XX AAD34322;

XX 16-JUL-2002 (first entry)

XX Human bone strength and mineralisation regulatory protein (BSMR) cDNA.
XX Human; bone strength and mineralisation regulatory protein; BSMR;
XX bone strength; mineralisation; ophthalmological; antidiabetic;
XX bone density regulating transmembrane receptor; prosthetic device;
XX surgical implant; diabetic retinopathy; hypertensive retinopathy;
XX therapy; osteoporosis; prematurity; ocular vessel; eye disorder;
XX osteopathic; gene; ss.
XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 52..4899

XX FT /*tag= a

XX FT /product= "Human BSMR protein"

XX WO200216553-A2.

XX 28-FEB-2002.

XX 17-AUG-2001; 2001WO-US041788.

XX 18-AUG-2000; 2000US-0226119P.

XX 22-SEP-2000; 2000US-0234337P.

XX 13-JUL-2001; 2001US-0304851P.

XX (AVET) AVENTIS PHARMA SA.

XX PA (HARD) HARVARD COLLEGE.

XX PA (UYCA-) UNIV CASE WESTERN RESERVE.

XX Warman ML, Gong Y, Olsen BR, Rawadi G, Roman-Roman S;

XX WPI; 2002-329694/36.

XX P-ESDB; AAE21730.

XX

QY	1801	CAGCTGCCCGACCTGATGGGGCTCAAAGCTGTGAATGTGGCAAGGTGCTCGGAACCAAC	1860
DB	1801	CAGCTGCCCGACCTGATGGGGCTCAAAGCTGTGAATGTGGCAAGGTGCTCGGAACCAAC	1860
QY	1861	CCGTGTGCGGACAGAAACGGGGGGTGACGACCACTGTGCTTTCTTCAACCCCAACCAAC	1920
DB	1861	CCGTGTGCGGACAGAAACGGGGGGTGACGACCACTGTGCTTTCTTCAACCCCAACCAAC	1920
QY	1921	CGGTGTGGCTGCCCATCGGCTCGAGCTGTGAGTGACATGAGACCTGTGCTGCT	1980
DB	1921	CGGTGTGGCTGCCCATCGGCTCGAGCTGTGAGTGACATGAGACCTGTGCTGCT	1980
QY	1981	GAGGCTTTCTTGGTCTTCAACAGAGAGCGCGCATCCACAGAGATCTCCCTCGAGACCAAT	2040
DB	1981	GAGGCTTTCTTGGTCTTCAACAGAGAGCGCGCATCCACAGAGATCTCCCTCGAGACCAAT	2040
QY	2041	AACAAACGAGCTCGGTGGAGCACTGCTGAGTGTGAGTTGGCTTTGACTACCCCGAGGGCATG	2100
DB	2041	AACAAACGAGCTCGGTGGAGCACTGCTGAGTGTGAGTTGGCTTTGACTACCCCGAGGGCATG	2100
QY	2101	GTGTCCCAACACCATCTACTGACAGACCTGAGCTGAGACCATGAGCGCGGCTTC	2160
DB	2101	GTGTCCCAACACCATCTACTGACAGACCTGAGCTGAGACCATGAGCGCGGCTTC	2160
QY	2161	ATGAAACGAGCTCGGTGGAGCACTGCTGAGTGTGAGTTGGCTTTGACTACCCCGAGGGCATG	2220
DB	2161	ATGAAACGAGCTCGGTGGAGCACTGCTGAGTGTGAGTTGGCTTTGACTACCCCGAGGGCATG	2220
QY	2221	GCCGTGTGACTGGATGGGCAAGAACTCTTACTGAGCGGACCACTGGGACCAACAGAAATCGAA	2280
DB	2221	GCCGTGTGACTGGATGGGCAAGAACTCTTACTGAGCGGACCACTGGGACCAACAGAAATCGAA	2280
QY	2281	GTGGCGGCTGAGCGGCACTGCGGCAAGTCTCGGCAAGTCTCGTGTGGAGGACTTGGACCAACCGG	2340
DB	2281	GTGGCGGCTGAGCGGCACTGCGGCAAGTCTCGGCAAGTCTCGTGTGGAGGACTTGGACCAACCGG	2340
QY	2341	AGTGTGCTGGGCTCGGATGCCCAACAGGGCTTACTGAGCGGAGCGGCTGAGGCGGCAAG	2400
DB	2341	AGTGTGCTGGGCTCGGATGCCCAACAGGGCTTACTGAGCGGAGCGGCTGAGGCGGCAAG	2400
QY	2401	CCGAGGATCTGTCGGGCTTTCATGGAACGGGACCAACTGCTGAGCGCTGCTGAGCAAGGTG	2460
DB	2401	CCGAGGATCTGTCGGGCTTTCATGGAACGGGACCAACTGCTGAGCGCTGCTGAGCAAGGTG	2460
QY	2461	GGCGGGGCAACGACCTCACCATTTGACTACGCTGACGAGCGGCTTACTGAGCAACCGG	2520
DB	2461	GGCGGGGCAACGACCTCACCATTTGACTACGCTGACGAGCGGCTTACTGAGCAACCGG	2520
QY	2521	GACACCAACATGATCGAGTCTGCAACATGCTGGGTGAGGAGCGGCTGCTGATTTGCCGAC	2580
DB	2521	GACACCAACATGATCGAGTCTGCAACATGCTGGGTGAGGAGCGGCTGCTGATTTGCCGAC	2580
QY	2581	GATCTCCGCAACCGCTTGGGTCTGACGAGTACGAGGATTTATCTACTGAGACAGCTGG	2640
DB	2581	GATCTCCGCAACCGCTTGGGTCTGACGAGTACGAGGATTTATCTACTGAGACAGCTGG	2640
QY	2641	AATCTGACACATGAGCGGGGCGGACAGACTAGCGGCGGAAACCGCACCTCATCCAG	2700
DB	2641	AATCTGACACATGAGCGGGGCGGACAGACTAGCGGCGGAAACCGCACCTCATCCAG	2700
QY	2701	GGCCACCTGGAATCTGATGAGACATCTGCTGCTTCCACTCTCCCGCAGGATGGGCTC	2760
DB	2701	GGCCACCTGGAATCTGATGAGACATCTGCTGCTTCCACTCTCCCGCAGGATGGGCTC	2760
QY	2761	AATGATGTATGCAACAAACGGGCGAGTGTGGGAGCTGTGCTTGGCATCCCGGGGCGG	2820
DB	2761	AATGATGTATGCAACAAACGGGCGAGTGTGGGAGCTGTGCTTGGCATCCCGGGGCGG	2820
QY	2821	CACGCTGGGCTGGGCTTCACTACACCTTGAACCCGACGAGCGGCAATGAGCGCGG	2880
DB	2821	CACGCTGGGCTGGGCTTCACTACACCTTGAACCCGACGAGCGGCAATGAGCGCGG	2880

QY	2881	CCCAACACCTTTCTGTCTTTCAGCCAGAAATCTGCCATCAGTCTCGATGATCCCGGACGAC	2940
DB	2881	CCCAACACCTTTCTGTCTTTCAGCCAGAAATCTGCCATCAGTCTCGATGATCCCGGACGAC	2940
QY	2941	CAGCACAGCCCGGATCTCATCTCTGCCCTGTGATGGAATGAGAAAGCTCAAAAGCCATCGAC	3000
DB	2941	CAGCACAGCCCGGATCTCATCTCTGCCCTGTGATGGAATGAGAAAGCTCAAAAGCCATCGAC	3000
QY	3001	TATGACCCACTGGAACAAGTTTCATCTACTGAGTGTGATGGGCGCGAGAACATCAAGGAGCC	3060
DB	3001	TATGACCCACTGGAACAAGTTTCATCTACTGAGTGTGATGGGCGCGAGAACATCAAGGAGCC	3060
QY	3061	AAGGACGACGAGCCAGCCCTTTGTTTGTGACCTCTCTGAGCCAAAGGCGGCAAGCCAGAC	3120
DB	3061	AAGGACGACGAGCCAGCCCTTTGTTTGTGACCTCTCTGAGCCAAAGGCGGCAAGCCAGAC	3120
QY	3121	AGGAGAGCCCAAGACCTCAGCATCGAATCTTACAGCCGGAACATCTGTTCTGGAAGTGGAG	3180
DB	3121	AGGAGAGCCCAAGACCTCAGCATCGAATCTTACAGCCGGAACATCTGTTCTGGAAGTGGAG	3180
QY	3181	GCCACCAATACCATCAACGTCACAGGCTGAGCGGGAAGCCATGGGGGTGGTGTCTGGT	3240
DB	3181	GCCACCAATACCATCAACGTCACAGGCTGAGCGGGAAGCCATGGGGGTGGTGTCTGGT	3240
QY	3241	GGGAGCCGCGAACAAGCCAGGCGCATCGTCTCAACGCGAGCGAGGCTACTGTTACTTC	3300
DB	3241	GGGAGCCGCGAACAAGCCAGGCGCATCGTCTCAACGCGAGCGAGGCTACTGTTACTTC	3300
QY	3301	ACCAACATGACGAGACCGGCGAGCCAAAGATCGAAACGCGAGCCCTGGAGCGGACCGAGCGC	3360
DB	3301	ACCAACATGACGAGACCGGCGAGCCAAAGATCGAAACGCGAGCCCTGGAGCGGACCGAGCGC	3360
QY	3361	GAGGTCTCTTTCACCAACCGGCTCATCGGCTCTGAGAGCTTGAAGCTGTGACGCTGTGAGG	3420
DB	3361	GAGGTCTCTTTCACCAACCGGCTCATCGGCTCTGAGAGCTTGAAGCTGTGACGCTGTGAGG	3420
QY	3421	GGCAAGCTGTCTGGGTGGAGCGGACCTGAAAGCGCATTTGAGAGCTGTGACGCTGTGAGG	3480
DB	3421	GGCAAGCTGTCTGGGTGGAGCGGACCTGAAAGCGCATTTGAGAGCTGTGACGCTGTGAGG	3480
QY	3481	GCCAAACCGCTTGAACCTTGGAGGAGCGGCAACATCGTGTGAGCTCTGAGGCTGTGACATCTT	3540
DB	3481	GCCAAACCGCTTGAACCTTGGAGGAGCGGCAACATCGTGTGAGCTCTGAGGCTGTGACATCTT	3540
QY	3541	GGCAAGCATCTCTACTGATCGAAGCGGCGGAGATGATCGAGCGGTGTGAGGAGGAC	3600
DB	3541	GGCAAGCATCTCTACTGATCGAAGCGGCGGAGATGATCGAGCGGTGTGAGGAGGAC	3600
QY	3601	ACCGGGGACAAAGCGGACCTCGCATCCAGGCGGCTGTGCGCCACCTCACTGCGCATCCATGCA	3660
DB	3601	ACCGGGGACAAAGCGGACCTCGCATCCAGGCGGCTGTGCGCCACCTCACTGCGCATCCATGCA	3660
QY	3661	GTGAGGAAAGTCAAGCTTGGAGGAGTCTCAGCCCAACCATGTGCCCGTGAACAATGGTGGC	3720
DB	3661	GTGAGGAAAGTCAAGCTTGGAGGAGTCTCAGCCCAACCATGTGCCCGTGAACAATGGTGGC	3720
QY	3721	TGCTCCCACTCTGATTTGCCAAGGCTGAGGAGACACAGGCTGTCTCATGCCAGTCCAC	3780
DB	3721	TGCTCCCACTCTGATTTGCCAAGGCTGAGGAGACACAGGCTGTCTCATGCCAGTCCAC	3780
QY	3781	CTCGTGTCTCTGCAAGAACCTGCTGACCTGTGAGAGCGGCGGCACTGTCTCCCGGACGAC	3840
DB	3781	CTCGTGTCTCTGCAAGAACCTGCTGACCTGTGAGAGCGGCGGCACTGTCTCCCGGACGAC	3840
QY	3841	TTTGATGTGTGCAACAGGGGAGATCGAATGTATCCCGGGGCTTGGCGCTGTGAAGGCTTTT	3900
DB	3841	TTTGATGTGTGCAACAGGGGAGATCGAATGTATCCCGGGGCTTGGCGCTGTGAAGGCTTTT	3900
QY	3901	CCCGAGTGTGATGACAGGAGGAGGCTTGGCGCTGTCTCCCGCGGCGGAGTTC	3960
DB	3901	CCCGAGTGTGATGACAGGAGGAGGCTTGGCGCTGTCTCCCGCGGCGGAGTTC	3960
QY	3961	CCCTGCGCGGGGTGAGTGTGGAACCTGCGGCTGCGGCTGCGAGCGGAGGAGCTGT	4020

[illegible]

Db 5041 GGCAGGGCTGGGAGACCTTTGTA 5063

RESULT 2

AAD34405

ID AAD34405 standard; cDNA; 5063 BP.

XX AC AAD34405;

XX 16-JUL-2002 (first entry)

XX Human BSMR mutant cDNA, G4416T.

XX Human; bone strength and mineralisation regulatory protein; BSMR;

KW bone strength; mineralisation; ophthalmological; antidiabetic;

KW bone density regulating transmembrane receptor; prosthetic device;

KW surgical implant; diabetic retinopathy; hypertensive retinopathy;

KW therapy; osteoporosis; prematurity; ocular vessel; eye disorder;

KW osteopathic; mutant; ss.

XX OS Homo sapiens.

OS Synthetic.

XX Key Location/Qualifiers

FT mutation replace(4416, G)

FT /*tag= a

XX W0200216553-A2.

XX 28-FEB-2002.

XX 17-AUG-2001; 2001WO-US041788.

XX 18-AUG-2000; 2000US-0226119P.

PR 22-SEP-2000; 2000US-0234337P.

PR 13-JUL-2001; 2001US-0304851P.

XX (AVET) AVENTIS PHARMA SA.

PA (HARD) HARVARD COLLEGE.

PA (UYCA-) UNIV CASE WESTERN RESERVE.

XX Warman ML, Gong Y, Olsen BR, Rawadi G, Roman-Roman S;

XX WPI; 2002-329694/36.

XX Polynucleotide encoding bone strength and mineralization regulatory

PT protein useful for diagnosis or therapy of osteoporosis.

XX Disclosure; Page; 124pp; English.

XX The invention relates to bone strength and mineralisation regulatory

CC protein (BSMR) and its corresponding nucleic acid sequence. BSMR DNA is

CC useful for the diagnosis or therapy of osteoporosis and for regulating

CC (increasing) bone strength and mineralisation in a human subject by

CC activating a bone density regulating transmembrane receptor (BSMR

CC protein). An expression vector comprising a promoter that is operably

CC linked to BSMR DNA is useful for modulating bone density and for

CC enhancing bone strength and mineralisation in a mammal cell. Composition

CC comprising a BSMR effector is useful for treating osteoporosis and is

CC useful particularly as a coating for prosthetic devices and surgical

CC implants. BSMR is useful for screening lead pharmaceutical agents as BSMR

CC effectors, which may be used to treat a range of eye disorders such as

CC diabetic retinopathy, hypertensive retinopathy and retinopathy of

CC prematurity, in which normal vascular growth and integrity of ocular

CC vessels is disrupted. The present sequence is human BSMR mutant cDNA.

CC Note: This sequence is not shown in the specification, however it is

CC constructed based on the human BSMR cDNA shown in fig 2 of the

CC specification (AAD34322)

XX Sequence 5063 BP; 1016 A; 1645 C; 1538 G; 864 T; 0 U; 0 Other;

SQ

Query Match 100.0%; Score 5061.4; DB 6; Length 5063;

Best Local Similarity 100.0%; Pred. No. 0;

QY 2161 ATGAACGGAGCTCGGTGGAGCAGTGTGTGAGTTTGGCCCTTGACTACCCCGAGGGCATG 2220
DB 2161 ATGAACGGAGCTCGGTGGAGCAGTGTGTGAGTTTGGCCCTTGACTACCCCGAGGGCATG 2220
QY 2221 GCCGTGTGACTGGGCAAGAACTCTTACTTGGCCGACACTGGGACCAACAGAAATCGAA 2280
DB 2221 GCCGTGTGACTGGGCAAGAACTCTTACTTGGCCGACACTGGGACCAACAGAAATCGAA 2280
QY 2281 GTGGCGGGCTGGACGGCGCAGTTCCGGCAAGTCTCTGTGTGGAGGGAATTGGACAACCCG 2340
DB 2281 GTGGCGGGCTGGACGGCGCAGTTCCGGCAAGTCTCTGTGTGGAGGGAATTGGACAACCCG 2340
QY 2341 AGGTCTGGCTGGCTGGATCCCAACCAAGGGCTACATCTACTGACCGAGTGGGGGGGCAAG 2400
DB 2341 AGGTCTGGCTGGCTGGATCCCAACCAAGGGCTACATCTACTGACCGAGTGGGGGGGCAAG 2400
QY 2401 CCGAGGATCTGGCGGGCTTCTATGACGGGACCAACTGCATGACGCTGTGACAAAGTG 2460
DB 2401 CCGAGGATCTGGCGGGCTTCTATGACGGGACCAACTGCATGACGCTGTGACAAAGTG 2460
QY 2461 GGCCGGGGCAACGAACTCAACATTGACTACGCTGACCGGCTCTTACTGGACCGACCTG 2520
DB 2461 GGCCGGGGCAACGAACTCAACATTGACTACGCTGACCGGCTCTTACTGGACCGACCTG 2520
QY 2521 GACACCAACATGATCGAGTGTCTCAACATGCTGGGTGAGGAGCGGTCGTGATTTGCCGAC 2580
DB 2521 GACACCAACATGATCGAGTGTCTCAACATGCTGGGTGAGGAGCGGTCGTGATTTGCCGAC 2580
QY 2581 GATCTCCGCAACCCGTTGGTCTGACGAGTACAGGATTAATCTACTGGACAGACTGG 2640
DB 2581 GATCTCCGCAACCCGTTGGTCTGACGAGTACAGGATTAATCTACTGGACAGACTGG 2640
QY 2641 AATCTGCACAGCAATTGAGCGGCGGCAAGAACTAGCGCGGAAACCGAACCTCATCCAG 2700
DB 2641 AATCTGCACAGCAATTGAGCGGCGGCAAGAACTAGCGCGGAAACCGAACCTCATCCAG 2700
QY 2701 GGCACCTGGACTTGTGTGATGACATCTCTGTGTTCCTCTCCGCGCAGGATGGCCTC 2760
DB 2701 GGCACCTGGACTTGTGTGATGACATCTCTGTGTTCCTCTCCGCGCAGGATGGCCTC 2760
QY 2761 AATGACTGTATGCACAAACGGGCGAGTGTGGCGAGCTGTGCTTGGCAATCCCGCGGC 2820
DB 2761 AATGACTGTATGCACAAACGGGCGAGTGTGGCGAGCTGTGCTTGGCAATCCCGCGGC 2820
QY 2821 CACCGTGGCTGGCTCTCACTTCCCTGCGATGGA CTGAGGAACGTCAAAGCCATCGAC 3000
DB 2821 CACCGTGGCTGGCTCTCACTTCCCTGCGATGGA CTGAGGAACGTCAAAGCCATCGAC 3000
QY 2881 CCCACCACTTCTGTGTTCAGCCAGAAATCTGCCATCAGTCGGATGATCCGGACGAC 2940
DB 2881 CCCACCACTTCTGTGTTCAGCCAGAAATCTGCCATCAGTCGGATGATCCGGACGAC 2940
QY 2941 CAGCAGACCCCGATCTCATCTCTGCGCCCTGCGATGGA CTGAGGAACGTCAAAGCCATCGAC 3000
DB 2941 CAGCAGACCCCGATCTCATCTCTGCGCCCTGCGATGGA CTGAGGAACGTCAAAGCCATCGAC 3000
QY 3001 TATGACCACTGGAAGTTTCTACTGTGGTGGATGGGGCCAGAACTCAAGCGAGCC 3060
DB 3001 TATGACCACTGGAAGTTTCTACTGTGGTGGATGGGGCCAGAACTCAAGCGAGCC 3060
QY 3061 AAGGACGAGGGACCCAGCCCTTGTGTTGACCTCTGAGCCAGGCGCAAAACCCAGAC 3120
DB 3061 AAGGACGAGGGACCCAGCCCTTGTGTTGACCTCTGAGCCAGGCGCAAAACCCAGAC 3120
QY 3121 AAGCAGCCCAACGACCTCAGACTCGACATCTACAGCCGACACTGTCTTGGAGCTGGAG 3180
DB 3121 AAGCAGCCCAACGACCTCAGACTCGACATCTACAGCCGACACTGTCTTGGAGCTGGAG 3180
QY 3181 GCCACCAATCACTCAAGTCTCAAGCTGAGCGGGAGCCATGGGGTGTGCTGCT 3240
DB 3181 GCCACCAATCACTCAAGTCTCAAGCTGAGCGGGAGCCATGGGGTGTGCTGCT 3240
QY 3241 GGGGACCGGCAAGCCCGAGGGCCATCTGTCTCAACGCGGAGCGGGTACTGTACTTC 3300

DB 3341 GGGGACCGGCAAGCCCGAGGGCCATCTGTCTCAACGCGGAGCGAGGTACTGTACTTC 3300
QY 3301 ACCAACTATGAGGACCGGGCAGCAAGATCGAAACGCGACGCTTGGACCGGACCGAGCGC 3360
DB 3301 ACCAACTATGAGGACCGGGCAGCAAGATCGAAACGCGACGCTTGGACCGGACCGAGCGC 3360
QY 3361 GAGTCTCTTACCAACCGGCTCATCCGCTGTGGCCCTGTGGTGGGACCAACACTG 3420
DB 3361 GAGTCTCTTACCAACCGGCTCATCCGCTGTGGCCCTGTGGTGGGACCAACACTG 3420
QY 3421 GGCAAGCTGTCTTGGGTGGACCGGCACTGAAAGCGCAATTGAGAGCTGTGACCTTCAAGG 3480
DB 3421 GGCAAGCTGTCTTGGGTGGACCGGCACTGAAAGCGCAATTGAGAGCTGTGACCTTCAAGG 3480
QY 3481 GCCAACCGCTGACCTTGGAGGACGCAACATCTGTGAGCCTCTGGGCTGTGACCATCTT 3540
DB 3481 GCCAACCGCTGACCTTGGAGGACGCAACATCTGTGAGCCTCTGGGCTGTGACCATCTT 3540
QY 3541 GGCAAGCTGTCTTACTGATCGACCGGCAAGCAGATGATCGAGCTGTGAGAGAGACC 3600
DB 3541 GGCAAGCTGTCTTACTGATCGACCGGCAAGCAGATGATCGAGCTGTGAGAGAGACC 3600
QY 3601 ACCGGGACAAAGCGGACTCTGATCCAGGCGCTGTGCGCCACCTCACTTGGCATCAATGCA 3660
DB 3601 ACCGGGACAAAGCGGACTCTGATCCAGGCGCTGTGCGCCACCTCACTTGGCATCAATGCA 3660
QY 3661 GTGAGGAAGTCACTGTGAGGAGTTCTCAGCCCAACCATGTGCGCTGTGACAAATGGTGGC 3720
DB 3661 GTGAGGAAGTCACTGTGAGGAGTTCTCAGCCCAACCATGTGCGCTGTGACAAATGGTGGC 3720
QY 3721 TGCTCCCACTCTTATTTGCAAGGAGTGGGACACCAAGCTGTCTATGCGCCAGTCCAC 3780
DB 3721 TGCTCCCACTCTTATTTGCAAGGAGTGGGACACCAAGCTGTCTATGCGCCAGTCCAC 3780
QY 3781 CTCTGTCTCTTGCAGAACCTCTGCTGACCTGTGAGAGCGCGCCCACTGTCTCCCGGACGAC 3840
DB 3781 CTCTGTCTCTTGCAGAACCTCTGCTGACCTGTGAGAGCGCGCCCACTGTCTCCCGGACGAC 3840
QY 3841 TTTGATGTGACAGGGGAGATCGACTGTATCTCCCGGGCTGTGCGCTGTGACGGCTTT 3900
DB 3841 TTTGATGTGACAGGGGAGATCGACTGTATCTCCCGGGCTGTGCGCTGTGACGGCTTT 3900
QY 3901 CCGAGTCCGATGACAGAGCGAGAGGCGCTGCGCTGTGCTGCTCCCGCGCCAGTTC 3960
DB 3901 CCGAGTCCGATGACAGAGCGAGAGGCGCTGCGCTGTGCTGCTCCCGCGCCAGTTC 3960
QY 3961 CCTGCGCGCGGGTCACTGTGTGACCTGTGCGCTGTGCGCTGTGCGAGCGGAGGACGTGT 4020
DB 3961 CCTGCGCGCGGGTCACTGTGTGACCTGTGCGCTGTGCGCTGTGCGAGCGGAGGACGTGT 4020
QY 4021 CAGGACCGCTCAGAGCGGCGGACTGTGACCGCACTCTGCTGCGCCCAACAGTTCGGTGT 4080
DB 4021 CAGGACCGCTCAGAGCGGCGGACTGTGACCGCACTCTGCTGCGCCCAACAGTTCGGTGT 4080
QY 4081 GCGAGCGGCGAGTGTCTCTATCAAAACAGCAGTGCAGCTCTTCCCGCAGTGTATCGAC 4140
DB 4081 GCGAGCGGCGAGTGTCTCTATCAAAACAGCAGTGCAGCTCTTCCCGCAGTGTATCGAC 4140
QY 4141 GGTCTCCGACGAGCTCATGTGTGAAATCAACAGCGCGCTCTCAGACGACGCGCGCCAC 4200
DB 4141 GGTCTCCGACGAGCTCATGTGTGAAATCAACAGCGCGCTCTCAGACGACGCGCGCCAC 4200
QY 4201 AGCAGTGCATCGGCGCGCTCATTTGGCATCATCTCTCTCTTGTGTGAGTGTGTGTC 4260
DB 4201 AGCAGTGCATCGGCGCGCTCATTTGGCATCATCTCTCTCTTGTGTGAGTGTGTGTC 4260
QY 4261 TATTTTGTGTGCGAGCGGTGTGTGCGAGGCTATGCGGGGGCCAAACGCGGCTTCCCG 4320
DB 4261 TATTTTGTGTGCGAGCGGTGTGTGCGAGGCTATGCGGGGGCCAAACGCGGCTTCCCG 4320
QY 4321 CACGAGTATGTGAGCGGAGCCCGCAGCTGCGCTCAATTTTCATAGCCCGCGGCTTC 4380

QY 361 ATCTCCGGCTGCTCTCTCCCGACGGGCTCGCTGCGAATGGGTGGGCAAGAGCTGTAC 420
DB 361 ATCTCCGGCTGCTCTCTCCCGACGGGCTCGCTGCGAATGGGTGGGCAAGAGCTGTAC 420
QY 421 TGGACGGACTCAGAGACAAACCGATCGAGTGGCCAACTCAATGGGCACATCCCGAAG 480
DB 421 TGGACGGACTCAGAGACAAACCGATCGAGTGGCCAACTCAATGGGCACATCCCGAAG 480
QY 481 GTGCTCTTCTGGCAGGACCTTGAACAGCTTAGGCCATTCGCTTGGACCCCGCTCAACGG 540
DB 481 GTGCTCTTCTGGCAGGACCTTGAACAGCTTAGGCCATTCGCTTGGACCCCGCTCAACGG 540
QY 541 TACATGTACTGGACAGACTGGGTGAGACGCCCCGGATTGAGCGGGCAGGATGGATGGC 600
DB 541 TACATGTACTGGACAGACTGGGTGAGACGCCCCGGATTGAGCGGGCAGGATGGATGGC 600
QY 601 AGCACCCGGAAGATCATTTGTGACTTCGACATTTTACTGGCCCAATGGACTGACCATCGAC 660
DB 601 AGCACCCGGAAGATCATTTGTGACTTCGACATTTTACTGGCCCAATGGACTGACCATCGAC 660
QY 661 CTGGAGGAGCAAGCTCTACTTGGCTGACCGCAAGCTCAGCTTCAATCCACCGTGCCAAAC 720
DB 661 CTGGAGGAGCAAGCTCTACTTGGCTGACCGCAAGCTCAGCTTCAATCCACCGTGCCAAAC 720
QY 721 CTGGACGGCTGTTCCGCGAGAGTGTGGAGGCGAGCTGACGACACCCCTTGGCCCTG 780
DB 721 CTGGACGGCTGTTCCGCGAGAGTGTGGAGGCGAGCTGACGACACCCCTTGGCCCTG 780
QY 781 ACGCTCTCCGGGACACTCTGTACTGACAGACTGGCAGACCCGCTCCATCCATCGCCCTGC 840
DB 781 ACGCTCTCCGGGACACTCTGTACTGACAGACTGGCAGACCCGCTCCATCCATCGCCCTGC 840
QY 841 AACAGCGCACTGGGGGGAAGAGAGATCTCTGAGTGGCTTCTACTCAACCCATGGAC 900
DB 841 AACAGCGCACTGGGGGGAAGAGAGATCTCTGAGTGGCTTCTACTCAACCCATGGAC 900
QY 901 ATCCAGGTGCTGAGCAGAGCGGCGAGCTTCTTCCACACTCGCTGTGAGGAGCAAT 960
DB 901 ATCCAGGTGCTGAGCAGAGCGGCGAGCTTCTTCCACACTCGCTGTGAGGAGCAAT 960
QY 961 GCGGCTGCTCCACTGTGCTGCTGCTCCCAAGCGAGCCTTCTACACATGCGCCCTGC 1020
DB 961 GCGGCTGCTCCACTGTGCTGCTGCTCCCAAGCGAGCCTTCTACACATGCGCCCTGC 1020
QY 1021 CCCACGGGTGTGACTGACAGCAAAACGCGAGACGTGTAAAGGACAGGACCGGAGGTG 1080
DB 1021 CCCACGGGTGTGACTGACAGCAAAACGCGAGACGTGTAAAGGACAGGACCGGAGGTG 1080
QY 1081 CTGCTGCTGGCCGCGGACGACCTACGGAGATCTCGCTGGACACGCGGACTTCACC 1140
DB 1081 CTGCTGCTGGCCGCGGACGACCTACGGAGATCTCGCTGGACACGCGGACTTCACC 1140
QY 1141 GACATCGTGTGCTGACAGTGAACATCCGGCAGCGCATTTGCCATCGACTAGACCCGCTA 1200
DB 1141 GACATCGTGTGCTGACAGTGAACATCCGGCAGCGCATTTGCCATCGACTAGACCCGCTA 1200
QY 1201 GAGGCTATGTCTATGGAAGATGACGAGTGGCGGCAATCCGACGGCGTACCTGGAC 1260
DB 1201 GAGGCTATGTCTATGGAAGATGACGAGTGGCGGCAATCCGACGGCGTACCTGGAC 1260
QY 1261 GGGTCTGGGCGCAGACCTGTGTCAACCGAGATCAACGACCCGATGGCATCCGGTTC 1320
DB 1261 GGGTCTGGGCGCAGACCTGTGTCAACCGAGATCAACGACCCGATGGCATCCGGTTC 1320
QY 1321 GACTGGGTGGCCCGAAACCTCTACTGGACCGACAGCGGACGACCGCATCGAGTGAAG 1380
DB 1321 GACTGGGTGGCCCGAAACCTCTACTGGACCGACAGCGGACGACCGCATCGAGTGAAG 1380
QY 1381 CCGCTCAACGGCACTCCCGCAAGATCTGTGTGTGGAGGACCTGACGAGCCCGAGCC 1440
DB 1381 CCGCTCAACGGCACTCCCGCAAGATCTGTGTGTGGAGGACCTGACGAGCCCGAGCC 1440

QY 1441 ATCGCACTGCACCCCGTATGGGCTCATGTACTGGACAGACTGGGAGAGAAACCTTAAA 1500
DB 1441 ATCGCACTGCACCCCGTATGGGCTCATGTACTGGACAGACTGGGAGAGAAACCTTAAA 1500
QY 1501 ATCAGAGTGTGCACAACTTTGGATGGGACGCGGTGTGTGGTCAATGCTCCTCGG 1560
DB 1501 ATCAGAGTGTGCACAACTTTGGATGGGACGCGCGTGTGTGGTCAATGCTCCTCGG 1560
QY 1561 TGGCCCAACGGCTGGCCCTGGACCTTCAGAGGGGAGTCTTACTGGGAGAGCCCAAG 1620
DB 1561 TGGCCCAACGGCTGGCCCTGGACCTTCAGAGGGGAGTCTTACTGGGAGAGCCCAAG 1620
QY 1621 ACAGCAAGATCGAGGTGATCAATGTTGATGGGACGAGGCGGACCTCCTCGAGGAC 1680
DB 1621 ACAGCAAGATCGAGGTGATCAATGTTGATGGGACGAGGCGGACCTCCTCGAGGAC 1680
QY 1681 AAGCTCCCGCACATTTTTCGGGTTCACTGCTGGGGGACTTCACTACTGAGTGAATG 1740
DB 1681 AAGCTCCCGCACATTTTTCGGGTTCACTGCTGGGGGACTTCACTACTGAGTGAATG 1740
QY 1741 CAGCGCCGACGATCGAGCGGGTGCACAAGTCAAGGCCAGCGCGGAGTCACTATTGAC 1800
DB 1741 CAGCGCCGACGATCGAGCGGGTGCACAAGTCAAGGCCAGCGCGGAGTCACTATTGAC 1800
QY 1801 CAGCTGCCCGACCTGATGGGCTCAAAAGCTGTGAATGTGGCCAAAGGTCTCGGAAACAA 1860
DB 1801 CAGCTGCCCGACCTGATGGGCTCAAAAGCTGTGAATGTGGCCAAAGGTCTCGGAAACAA 1860
QY 1861 CCGTGTGCGGACAGAAACGGGGGTGAGCCACTGCTGCTTCTTCAACCCACGCAAC 1920
DB 1861 CCGTGTGCGGACAGAAACGGGGGTGAGCCACTGCTGCTTCTTCAACCCACGCAAC 1920
QY 1921 CGTGTGGCTGCCCATCGGCTGAGCTGTGATGATGATGATGATGATGATGATGATGAT 1980
DB 1921 CGTGTGGCTGCCCATCGGCTGAGCTGTGATGATGATGATGATGATGATGATGATGAT 1980
QY 1981 GAGGCTTCTTGTGCTTCCACGACAGAGCGCCCATCCACAGGATCTCCCTCGAGACCAAT 2040
DB 1981 GAGGCTTCTTGTGCTTCCACGACAGAGCGCCCATCCACAGGATCTCCCTCGAGACCAAT 2040
QY 2041 AACAAAGAGTGGCCATCCCGCTCAAGGCGCTCAAGGAGGCTCAGCCCTGAGCTTGTAT 2100
DB 2041 AACAAAGAGTGGCCATCCCGCTCAAGGCGCTCAAGGAGGCTCAGCCCTGAGCTTGTAT 2100
QY 2101 GTGTCCAAACACCATCTACTGACAGAGCTGAGAGCCATCAGCGCGGCTTC 2160
DB 2101 GTGTCCAAACACCATCTACTGACAGAGCTGAGAGCCATCAGCGCGGCTTC 2160
QY 2161 ATGAAACGGGAGCTCGGTGGAGCAGCTGTGAGTGTGGCTTGAATACCCCGAGGCGATG 2220
DB 2161 ATGAAACGGGAGCTCGGTGGAGCAGCTGTGAGTGTGGCTTGAATACCCCGAGGCGATG 2220
QY 2221 GCGGTGTGATGGGCAAGAACTCTACTGGGCGGACACTGGGACCAACAGATCGAA 2280
DB 2221 GCGGTGTGATGGGCAAGAACTCTACTGGGCGGACACTGGGACCAACAGATCGAA 2280
QY 2281 GTGGCGGCTGAGCGGCGAGTCCGCAAGTCTCTGTGTGGAGGACTTGGACAAACCG 2340
DB 2281 GTGGCGGCTGAGCGGCGAGTCCGCAAGTCTCTGTGTGGAGGACTTGGACAAACCG 2340
QY 2341 AGGTGCTGGCCCTGGATCCCAAGGCTTACATCTACTGGAACGAGTGGGCGGCAAG 2400
DB 2341 AGGTGCTGGCCCTGGATCCCAAGGCTTACATCTACTGGAACGAGTGGGCGGCAAG 2400
QY 2401 CCGAGGATCTGGGCGCTTCAATGGAACGAGCACTGATGAGCGCTGGTGGACAGGTG 2460
DB 2401 CCGAGGATCTGGGCGCTTCAATGGAACGAGCACTGATGAGCGCTGGTGGACAGGTG 2460
QY 2461 GCGCGGGCAACGACCTCAATTTGATCACTGACAGCGCTTACTTGAACCGACCTG 2520
DB 2461 GCGCGGGCAACGACCTCAATTTGATCACTGACAGCGCTTACTTGAACCGACCTG 2520
QY 2521 GACACCAACATGATCGAGTCTGTCACCAATGTGGGTTCAGGAGCGGCTGTGATTGCCGAC 2580

[illegible]

Db	3601	ACCGGGGACAAAGCGGACTCGCATCCAGGGCGGTGTGCGCCCACTCACTGGCATCCATGCA	3661
Qy	3661	GTGAGGAAGTCAAGCTGAGGAGTCTTCAAGCCCAACCATGTGCGGTGACAAATGTTGGC	3720
Db	3661	GTGAGGAAAGTCAAGCTGAGGAGTCTTCAAGCCCAACCATGTGCGGTGACAAATGTTGGC	3720
Qy	3721	TGCTCCCACATCTGTATTTGCCAAGGTGTATGGGACACCAAGGTGTCAATGCCCAGTCCAC	3780
Db	3721	TGCTCCCACATCTGTATTTGCCAAGGTGTATGGGACACCAAGGTGTCAATGCCCAGTCCAC	3780
Qy	3781	CTCGTGTCTTGCAAGACCTGCTGACCTGTGTGAGAGCGGCCCACTGCTCCCGGACCCAG	3840
Db	3781	CTCGTGTCTTGCAAGACCTGCTGACCTGTGTGAGAGCGGCCCACTGCTCCCGGACCCAG	3840
Qy	3841	TTTGCATGTGCCACAGGGGAGATCGACTGTATCCCCGGGGCTTGGCGCTGTGACGGCTTT	3900
Db	3841	TTTGCATGTGCCACAGGGGAGATCGACTGTATCCCCGGGGCTTGGCGCTGTGACGGCTTT	3900
Qy	3901	CCCAGTGTGCATGATCAGAGACGAGAGGGGTGCCCCGTGTGCTCCGCCGCCACAGTTC	3960
Db	3901	CCCAGTGTGCATGATCAGAGACGAGAGGGGTGCCCCGTGTGCTCCGCCGCCACAGTTC	3960
Qy	3961	CCCTGCGCGGGGTCACTGTGTGGACCTGCGCTTGGCTGCGACGGCGAGGACACTGT	4020
Db	3961	CCCTGCGCGGGGTCACTGTGTGGACCTGCGCTTGGCTGCGACGGCGAGGACACTGT	4020
Qy	4021	CAGACCGCTCAGACGAGGCGGACTGTGACGCCATCTGCCTGCCCAACAGTTCGGGTGT	4080
Db	4021	CAGACCGCTCAGACGAGGCGGACTGTGACGCCATCTGCCTGCCCAACAGTTCGGGTGT	4080
Qy	4081	GCGAGCGGCGAGTGTCTCATCAACACGAGTGGGACTCTTCCCGACTGTATGAC	4140
Db	4081	GCGAGCGGCGAGTGTCTCATCAACACGAGTGGGACTCTTCCCGACTGTATGAC	4140
Qy	4141	GGCTCCGACGAGTCACTGTGTGAAATCACCAAGCGGCTTCCAGACGAGCCCGGCCAC	4200
Db	4141	GGCTCCGACGAGTCACTGTGTGAAATCACCAAGCGGCTTCCAGACGAGCCCGGCCAC	4200
Qy	4201	AGCAGTGCATCGGGCCCGTCAATTGGCATCATCTCTCTCTTCTGTGTCATGGGTGTGTC	4260
Db	4201	AGCAGTGCATCGGGCCCGTCAATTGGCATCATCTCTCTCTTCTGTGTCATGGGTGTGTC	4260
Qy	4261	TATTTTGTGTGCCAGCGCGTGTGTCAGCGTATGCGGGGCCCAAGCGGCCCTTCCCG	4320
Db	4261	TATTTTGTGTGCCAGCGCGTGTGTCAGCGTATGCGGGGCCCAAGCGGCCCTTCCCG	4320
Qy	4321	CACAGTATGTACGCGGACCCGCGCAGTGCCTCAATTTCAATAGCCCGGGCGGTTC	4380
Db	4321	CACAGTATGTACGCGGACCCGCGCAGTGCCTCAATTTCAATAGCCCGGGCGGTTC	4380
Qy	4381	CAGCATGGCCCTTTCACAGGCATCGCATGGGAAAGTCCATGATGAGTCCGTGAGCCTG	4440
Db	4381	CAGCATGGCCCTTTCACAGGCATCGCATGGGAAAGTCCATGATGAGTCCGTGAGCCTG	4440
Qy	4441	ATGGGGGGCGGGCGGGGTGCTGTACGACCGGAAACCAAGTTCAGAGGGGCTCGTCC	4500
Db	4441	ATGGGGGGCGGGCGGGGTGCTGTACGACCGGAAACCAAGTTCAGAGGGGCTCGTCC	4500
Qy	4501	AGCAGTGTGTACGACGAAAGGCCAGCTGTATCCCGCGATTCGTAAACCCGCGCCCTCC	4560
Db	4501	AGCAGTGTGTACGACGAAAGGCCAGCTGTATCCCGCGATTCGTAAACCCGCGCCCTCC	4560
Qy	4561	CCGGCCACGACCCCTCCCTGTACACATGGACATGTTCTTCTTCAACATTCGGGCC	4620
Db	4561	CCGGCCACGACCCCTCCCTGTACACATGGACATGTTCTTCTTCAACATTCGGGCC	4620
Qy	4621	ACTGCGAGCCGTACAGGCCCTACATCTTCGAGGAATGGCGCCCGCCGACGCGCCTGC	4680
Db	4621	ACTGCGAGCCGTACAGGCCCTACATCTTCGAGGAATGGCGCCCGCCGACGCGCCTGC	4680
Qy	4681	AGCACCGAGTGTGTACAGGACTACAGCGCCAGCGCTGTGAAGGCGAGCAAGTACTAC	4740
Db	4681	AGCACCGAGTGTGTACAGGACTACAGCGCCAGCGCTGTGAAGGCGAGCAAGTACTAC	4740

QY 721 CTGACGGCTGTTCCGGCAGAGGTGTTGGAGGCGACCTGACGCAACCCCTTGCCCTGC 780
DB 721 CTGAGCGCTCGTTCCGGCAGAGGTGTTGGAGGCGACCTGACGCAACCCCTTGCCCTGC 780
QY 781 ACGCTCTCCGGGACACTCTGTACTGGACAGACTGGCAGACCCGCTCATCATGCTGTC 840
DB 781 ACGCTCTCCGGGACACTCTGTACTGGACAGACTGGCAGACCCGCTCATCATGCTGTC 840
QY 841 AACAGCGCACTGGGGGAAGAGGAGAGTCTGTAGTGCCTCTACTCAACCCATGGAC 900
DB 841 AACAGCGCACTGGGGGAAGAGGAGAGTCTGTAGTGCCTCTACTCAACCCATGGAC 900
QY 901 ATCCAGGTGCTGAGCCAGGACGCGAGCTTTCTTCCACACTCGCTGTGAGGAGACAT 960
DB 901 ATCCAGGTGCTGAGCCAGGACGCGAGCTTTCTTCCACACTCGCTGTGAGGAGACAT 960
QY 961 GCGGCTGCTCCCACTGTGCTGTGTTGCCCAAGCGAGGCTTTCTACATGCGCTGTC 1020
DB 961 GCGGCTGCTCCCACTGTGCTGTGTTGCCCAAGCGAGGCTTTCTACATGCGCTGTC 1020
QY 1021 CCCACGGTGTGACGTGCAGACAAACGGCAGGAGCTGTAGGCGAGGACCGGAGGAGTG 1080
DB 1021 CCCACGGTGTGACGTGCAGACAAACGGCAGGAGCTGTAGGCGAGGACCGGAGGAGTG 1080
QY 1081 CTGCTGTGCTGCGCGGACGGAACCTACGAGGATCTCGCTGGACACGCCGGAATTCAAC 1140
DB 1081 CTGCTGTGCTGCGCGGACGGAACCTACGAGGATCTCGCTGGACACGCCGGAATTCAAC 1140
QY 1141 GACATGTGTGACAGGTGGAACGACATCCGGCAGCGCATTTGCCATGCACTACGACCCGCTA 1200
DB 1141 GACATGTGTGACAGGTGGAACGACATCCGGCAGCGCATTTGCCATGCACTACGACCCGCTA 1200
QY 1201 GAGGCTATGTCTACTGACAGATGACGAGGTGCGGCGCATCCGAGGCGGTACTCTGGAC 1260
DB 1201 GAGGCTATGTCTACTGACAGATGACGAGGTGCGGCGCATCCGAGGCGGTACTCTGGAC 1260
QY 1261 GGGTCTGGGGCGACGCTGCTCAACACCGAGATCAACGACCCCGATGGCATCGCGTC 1320
DB 1261 GGGTCTGGGGCGACGCTGCTCAACACCGAGATCAACGACCCCGATGGCATCGCGTC 1320
QY 1321 GACTGGGTGGCCGAACCTCTACTGACCGACACGGGCAACCGCATCGAGGTGACG 1380
DB 1321 GACTGGGTGGCCGAACCTCTACTGACCGACACGGGCAACCGCATCGAGGTGACG 1380
QY 1381 GCGCTCAACGCGACCTCCGCAAGATCTGTGTGCGAGGACCTGGACGAGCCCGGAGCC 1440
DB 1381 GCGCTCAACGCGACCTCCGCAAGATCTGTGTGCGAGGACCTGGACGAGCCCGGAGCC 1440
QY 1441 ATCGCACTGCACCCCGTGAATGGGCTCATGTACTGGACAGACTGGGAGAGAACCTTAA 1500
DB 1441 ATCGCACTGCACCCCGTGAATGGGCTCATGTACTGGACAGACTGGGAGAGAACCTTAA 1500
QY 1501 ATCGAGTGTGCAACTTTGGATGGGAGGAGCGGCTGTGTGTTCAATGCTCCTCGGG 1560
DB 1501 ATCGAGTGTGCAACTTTGGATGGGAGGAGCGGCTGTGTGTTCAATGCTCCTCGGG 1560
QY 1561 TGGCCCAACGCGCTGGCCCTGGAACCTGACAGAGGAGGAGCTCTACTGGGAGAGCGCAAG 1620
DB 1561 TGGCCCAACGCGCTGGCCCTGGAACCTGACAGAGGAGGAGCTCTACTGGGAGAGCGCAAG 1620
QY 1621 ACAGACAAGATCGAGGTGATCAATGTTGTATGGGACGAGGCGGACCCCTCTCGAGGAC 1680
DB 1621 ACAGACAAGATCGAGGTGATCAATGTTGTATGGGACGAGGCGGACCCCTCTCGAGGAC 1680
QY 1681 AAGCTCCGCAATTTTCGGGTTCAACGCTGTGCGGGGACTTCACTACTGAGTGAATG 1740
DB 1681 AAGCTCCGCAATTTTCGGGTTCAACGCTGTGCGGGGACTTCACTACTGAGTGAATG 1740
QY 1741 CAGCGCGCAGCATCGAGCGGTGCAAGAGTCAAGCCAGCGGAGCTCATATTGAC 1800
DB 1741 CAGCGCGCAGCATCGAGCGGTGCAAGAGTCAAGCCAGCGGAGCTCATATTGAC 1800
QY 1801 CAGCTGCCCGACCTGATGGGCTCAAGCTGTGAATGTGGCCAAAGTCTGCGAACAAC 1860

DB 1801 CAGCTGCCCGACCTGATGGGCTCAAGCTGTGAATGTGGCCAAAGTCTGTCGGAACCAAC 1860
QY 1861 CCGTGTGCGGACAGAAACGGGGGTGACGCCACCTGTGCTTCTTTCACACCCACGCAAC 1920
DB 1861 CCGTGTGCGGACAGAAACGGGGGTGACGCCACCTGTGCTTCTTTCACACCCACGCAAC 1920
QY 1921 CCGTGTGCTGCCCATCGGCTCGAGCTGTGAGTGACATGAAGACCTGATCGTGCT 1980
DB 1921 CCGTGTGCTGCCCATCGGCTCGAGCTGTGAGTGACATGAAGACCTGATCGTGCT 1980
QY 1981 GAGGCTTTCTTGGTCTTCCACAGCAGAGCCGCCATCCACAGGATCTCCTCTGAGACCAAT 2040
DB 1981 GAGGCTTTCTTGGTCTTCCACAGCAGAGCCGCCATCCACAGGATCTCCTCTGAGACCAAT 2040
QY 2041 AACAAACGAGTGGGCTCCCGCTCAGCGGCTCAAGAGGCTCAGCCCTGAGCTTTGAT 2100
DB 2041 AACAAACGAGTGGGCTCCCGCTCAGCGGCTCAAGAGGCTCAGCCCTGAGCTTTGAT 2100
QY 2101 GTGTCCAAACACACATCTACTGGACAGACGTCAAGCTTGAAGACCATCAGCCGCGCTTC 2160
DB 2101 GTGTCCAAACACACATCTACTGGACAGACGTCAAGCTTGAAGACCATCAGCCGCGCTTC 2160
QY 2161 ATGAACGGGAGCTCGGTGGACACGTGTGAGTGTGGCTTGTACTACCCGAGGCGATG 2220
DB 2161 ATGAACGGGAGCTCGGTGGACACGTGTGAGTGTGGCTTGTACTACCCGAGGCGATG 2220
QY 2221 GCGCTTGACTGGATGGGCAAGAACCTCTACTGGGCGGACACTGGGACCAACAGATCGAA 2280
DB 2221 GCGCTTGACTGGATGGGCAAGAACCTCTACTGGGCGGACACTGGGACCAACAGATCGAA 2280
QY 2281 GTGCGCGGCTGGACGGGAGTCCGCAAGTCTCGTGTGGAGGAGCTTGGACAAACCG 2340
DB 2281 GTGCGCGGCTGGACGGGAGTCCGCAAGTCTCGTGTGGAGGAGCTTGGACAAACCG 2340
QY 2341 AGGTGCTGCGCTGGCTTCCCAACAGGCTTACTACTGGACCGAGTGGGCGGCGAAG 2400
DB 2341 AGGTGCTGCGCTGGCTTCCCAACAGGCTTACTACTGGACCGAGTGGGCGGCGAAG 2400
QY 2401 CCGAGGATCGTGGGCTTTCATGACGGGACCAACTGTCATGACGCTGGTGGACAGGTG 2460
DB 2401 CCGAGGATCGTGGGCTTTCATGACGGGACCAACTGTCATGACGCTGGTGGACAGGTG 2460
QY 2461 GCGCGGCGCAACGACCTCACCATTGACTAGCTGACGAGCGGCTTACTGACCGACCTG 2520
DB 2461 GCGCGGCGCAACGACCTCACCATTGACTAGCTGACGAGCGGCTTACTGACCGACCTG 2520
QY 2521 GACACCAACATGATCGAGTCTGTCACCAATGCTGGGTGAGGAGCGGCTGATTTGCCGAC 2580
DB 2521 GACACCAACATGATCGAGTCTGTCACCAATGCTGGGTGAGGAGCGGCTGATTTGCCGAC 2580
QY 2581 GATCTCCGCAACCGCTTTCATGAGGAGTACAGGATTTATCTACTGACAGACTGG 2640
DB 2581 GATCTCCGCAACCGCTTTCATGAGGAGTACAGGATTTATCTACTGACAGACTGG 2640
QY 2641 AATCTGACAGCATTTGAGCGGCGCAAGACTAGCGGCGGAAACCGCACCTCATCCAG 2700
DB 2641 AATCTGACAGCATTTGAGCGGCGCAAGACTAGCGGCGGAAACCGCACCTCATCCAG 2700
QY 2701 GGGCACCTGAGCTTCTGTGATGGAATCTGCTGTGTTTCCACTCTCCCGCAGGATGGCTC 2760
DB 2701 GGGCACCTGAGCTTCTGTGATGGAATCTGCTGTGTTTCCACTCTCCCGCAGGATGGCTC 2760
QY 2761 AATGATGTATGCAACAAACGGGAGTGTGGGAGCTGTGCTTGCATCCCGGCGGCG 2820
DB 2761 AATGATGTATGCAACAAACGGGAGTGTGGGAGCTGTGCTTGCATCCCGGCGGCG 2820
QY 2821 CACGCTGCGGCTGCGCTCACACTACACCTTGACCCCGAGCGGCACTGAGCCGCG 2880
DB 2821 CACGCTGCGGCTGCGCTCACACTACACCTTGACCCCGAGCGGCACTGAGCCGCG 2880
QY 2881 CCCACCAACCTTCTTGTGTTTCCAGCCAGAAATCTGCCATCAGTCCGATGATCCCGAGCAG 2940

QY 1381 GCGCTCAACGCGACCTCCGCAAGATCTGTGTGCGAGGACCTGAGAGCCCGAGCC 1440
DB 1400 GCGCTCAACGCGACCTCCGCAAGATCTGTGTGCGAGGACCTGAGAGCCCGAGCC 1459
QY 1441 ATCGCACTGCAACCGGCTGATGGGCTCATGTACTGGACAGACTGGGGAGAGAACCTTAAA 1500
DB 1460 ATCGCACTGCAACCGGCTGATGGGCTCATGTACTGGACAGACTGGGGAGAGAACCTTAAA 1519
QY 1501 ATCGAGTGTGCAACTTGGATGGGAGAGAGCGGCTGTGCTGCTCAATGCCCTCCCTCGGG 1560
DB 1520 ATCGAGTGTGCAACTTGGATGGGAGAGAGCGGCTGTGCTGCTCAATGCCCTCCCTCGGG 1579
QY 1561 TGGCCCAACGCGCTCGGCTCGACCTGACAGAGGGAAGCTCTACTGGGGAGAGCCCAAG 1620
DB 1580 TGGCCCAACGCGCTCGGCTCGACCTGACAGAGGGAAGCTCTACTGGGGAGAGCCCAAG 1639
QY 1621 ACAGACAAGATCGAGGTGATCAATGTGATGGGACGAAGGCGGACCTCTCTGAGGAC 1680
DB 1640 ACAGACAAGATCGAGGTGATCAATGTGATGGGACGAAGGCGGACCTCTCTGAGGAC 1699
QY 1681 AAGCTCCGCAATTTTCGGGTTCAGGCTGTGCGGGGACTTCATCTACTGGAAGTACTGG 1740
DB 1700 AAGCTCCGCAATTTTCGGGTTCAGGCTGTGCGGGGACTTCATCTACTGGAAGTACTGG 1759
QY 1741 CAGCGCCGACGATCGAGCGGTGCAAAAGGTCAAGGCGAGCGGAGCGTCAATATTGAC 1800
DB 1760 CAGCGCCGACGATCGAGCGGTGCAAAAGGTCAAGGCGAGCGGAGCGTCAATATTGAC 1819
QY 1801 CAGCTGCCCGACCTGATGGGCTCAAAGTGTGAAATGTGGCCAAAGGTGTCGGAACCAAC 1860
DB 1820 CAGCTGCCCGACCTGATGGGCTCAAAGTGTGAAATGTGGCCAAAGGTGTCGGAACCAAC 1879
QY 1861 CCGTGTGCGGACAGAAACGGGGGTGACGCCACCTGTGCTGCTTTCACACCCGACGCAACC 1920
DB 1880 CCGTGTGCGGACAGAAACGGGGGTGACGCCACCTGTGCTGCTTTCACACCCGACGCAACC 1939
QY 1921 CCGTGTGCGCTGCCCATCGGCTGAGCTGTGAGTGAATGATGAAGACCTGTGATCGTCT 1980
DB 1940 CCGTGTGCGCTGCCCATCGGCTGAGCTGTGAGTGAATGATGAAGACCTGTGATCGTCT 1999
QY 1981 GAGGCTTCTTGGTCTTCAACAGCAGAGCGGCCATCCACAGGATCTCCCTCGAGACCAAT 2040
DB 2000 GAGGCTTCTTGGTCTTCAACAGCAGAGCGGCCATCCACAGGATCTCCCTCGAGACCAAT 2059
QY 2041 AACACGAGTGGCCATCCCGCTACGCGGCTCAAGAGGCTCAGCCCTGAGCTTTCAT 2100
DB 2060 AACACGAGTGGCCATCCCGCTACGCGGCTCAAGAGGCTCAGCCCTGAGCTTTCAT 2119
QY 2101 GTGTCCAAACACACATCTACTGGACAGACGTCAGCCTGAAGACCATCAGCCGCGCTTC 2160
DB 2120 GTGTCCAAACACACATCTACTGGACAGACGTCAGCCTGAAGACCATCAGCCGCGCTTC 2179
QY 2161 ATGAACGAGGAGTCCGTGGAGACGCTGTGTGAGTGTGGCTTTGACTAATCCCGAGGGCATG 2220
DB 2180 ATGAACGAGGAGTCCGTGGAGACGCTGTGTGAGTGTGGCTTTGACTAATCCCGAGGGCATG 2239
QY 2221 GCGTGTGATGAGTGGGCAAGAACCTTACTGGGCGGACACTGGGACCAACAGAAATCGAA 2280
DB 2240 GCGTGTGATGAGTGGGCAAGAACCTTACTGGGCGGACACTGGGACCAACAGAAATCGAA 2299
QY 2281 GTGGCGGCTGGACGGGCGAGTTCGGGCAAGTTCCTGTGTGAGGAGCTTGGACAAACCGG 2340
DB 2300 GTGGCGGCTGGACGGGCGAGTTCGGGCAAGTTCCTGTGTGAGGAGCTTGGACAAACCGG 2359
QY 2341 AGGTGCTGCGCTGGATCCCAACCAAGGCTACATCTACTGAGCCGAGTGGGGCGGCAAG 2400
DB 2360 AGGTGCTGCGCTGGATCCCAACCAAGGCTACATCTACTGAGCCGAGTGGGGCGGCAAG 2419
QY 2401 CCGAGGATGCTGGGCGCTTTCATGAGCGGACCAACTGATCACTGCTGTGAGCAAGGTG 2460
DB 2420 CCGAGGATGCTGGGCGCTTTCATGAGCGGACCAACTGATCACTGCTGTGAGCAAGGTG 2479
QY 2461 GGCGCGGCGCAACGACCTCACCATTGACTACGCTGACCGGCTCTACTGGACCGACCTG 2520

DB 2480 GGCGCGGCGCAACGACCTCACCATTGACTACGCTGACCGGCTCTACTGGACCGACCTG 2539
QY 2521 GACACCAACATGATCGAGTGTGCTCAACATGCTGGGTGAGGCGGGTGTGATTTGCCGAC 2580
DB 2540 GACACCAACATGATCGAGTGTGCTCAACATGCTGGGTGAGGCGGGTGTGATTTGCCGAC 2599
QY 2581 GATCTCCGCGACCGGCTGCTGACGAGTACGAGGATATATATCTACTGACAGACTGG 2640
DB 2600 GATCTCCGCGACCGGCTGCTGAGCGAGTACGAGGATATATATCTACTGACAGACTGG 2659
QY 2641 AATCTGACAGCAATTTGAGCGGGCGGACAAAGACTAGCGGCGGAAACCGCACCTCATCCAG 2700
DB 2660 AATCTGACAGCAATTTGAGCGGGCGGACAAAGACTAGCGGCGGAAACCGCACCTCATCCAG 2719
QY 2701 GGCGACCTTGTGATGAGGATCTCTGCTGTGCTCACTCTCCCGCAGGATGGCCTC 2760
DB 2720 GGCGACCTTGTGATGAGGATCTCTGCTGTGCTCACTCTCCCGCAGGATGGCCTC 2779
QY 2761 AATGACTGTATGACAAACCGGCGAGTGTGGGAGCTGTGCTTGCCTCCCGGCGGC 2820
DB 2780 AATGACTGTATGACAAACCGGCGAGTGTGGGAGCTGTGCTTGCCTCCCGGCGGC 2839
QY 2821 CACGCTGCGGCTGCGCTCACTACACCTTGAACCCAGCAGCCGCAACTGACGCGC 2880
DB 2840 CACGCTGCGGCTGCGCTCACTACACCTTGAACCCAGCAGCCGCAACTGACGCGC 2899
QY 2881 CCGACACCTTCTTGTGCTTTCAGCCAGAAATCTGCCATCAGTCGATGATCCCGGACGAC 2940
DB 2900 CCGACACCTTCTTGTGCTTTCAGCCAGAAATCTGCCATCAGTCGATGATCCCGGACGAC 2959
QY 2941 CAGCAGCGCGGATCTCATCTGCCCTGATGAGCTGAGGAAACGTCAGGACCAATCGAC 3000
DB 2960 CAGCAGCGCGGATCTCATCTGCCCTGATGAGCTGAGGAAACGTCAGGACCAATCGAC 3019
QY 3001 TATGACCCACTGGAACAGTTCATCTACTGGTGTGATGGGCGCCAGAAATCAAGCGAGCC 3060
DB 3020 TATGACCCACTGGAACAGTTCATCTACTGGTGTGATGGGCGCCAGAAATCAAGCGAGCC 3079
QY 3061 AAGGACGCGGAGCCAGCCCTTGTGCTGAGCTCTGAGCCAGGCGCAAAACCCAGAC 3120
DB 3080 AAGGACGCGGAGCCAGCCCTTGTGCTGAGCTCTGAGCCAGGCGCAAAACCCAGAC 3139
QY 3121 AGGACGCGCCACGACCTCAGCATCGACATCTACAGCCGGAACACTGTTCTGGACGTCGAG 3180
DB 3140 AGGACGCGCCACGACCTCAGCATCGACATCTACAGCCGGAACACTGTTCTGGACGTCGAG 3199
QY 3181 GCCACCAATACATCAACGCTCCAAGGCTGAGCGGGGAAGCCATGGGGGTGTGCTGCT 3240
DB 3200 GCCACCAATACATCAACGCTCCAAGGCTGAGCGGGGAAGCCATGGGGGTGTGCTGCT 3259
QY 3241 GGGGACGCGGACCAAGCCAGGGCCATCGTCTCAACGCGGAGCGGCTGCTGCTGCTTTC 3300
DB 3260 GGGGACGCGGACCAAGCCAGGGCCATCGTCTCAACGCGGAGCGGCTGCTGCTGCTTTC 3319
QY 3301 ACCAATGACGAGGACCGGGCAGCCAAAGATCGAAGCGCAGCCCTGGAACCGGACGCGC 3360
DB 3320 ACCAATGACGAGGACCGGGCAGCCAAAGATCGAAGCGCAGCCCTGGAACCGGACGCGC 3379
QY 3361 GAGGTCTCTTTCACACCGGCTCATTCGCGCTGTGCGCTGTGGTGTGGAACACACTG 3420
DB 3380 GAGGTCTCTTTCACACCGGCTCATTCGCGCTGTGCGCTGTGGTGTGGAACACACTG 3439
QY 3421 GCGAGCTGTCTGGGTGAGCGGACCTGAAGCGCATTTGAGAGCTGTGACTGTGAGG 3480
DB 3440 GCGAGCTGTCTGGGTGAGCGGACCTGAAGCGCATTTGAGAGCTGTGACTGTGAGG 3499
QY 3481 GCCAACCCCTGACCTCGGAGGACGCGCAACATCTGTGAGCCTCTGGGCTGACCAATCTT 3540
DB 3500 GCCAACCCCTGACCTCGGAGGACGCGCAACATCTGTGAGCCTCTGGGCTGACCAATCTT 3559
QY 3541 GCGAAGCATCTCTACTGATGACCGCGCAGCAGATGATCGAGCTGTGAGAGAGCC 3600

Db 3560 GCGAAGCATCTCTACTGGATCGACCGCCAGCAGCAGATGATCGAGCGTGTGAGAGACC 3619
Qy 3601 ACCGGGGACAAGCGGACTCGCATCCAGGCGCGTGTGCGCCCACTCACTCGGATCCATGCA 3660
Db 3620 ACCGGGGACAAGCGGACTCGCATCCAGGCGCGTGTGCGCCCACTCACTCGGATCCATGCA 3679
Qy 3661 GTGGAGGAAGTCAGCTCGGAGGAGTTCTCAGCCCAACCATGTGCCCGTGACAAATGGTGGC 3720
Db 3680 GTGGAGGAAGTCAGCTCGGAGGAGTTCTCAGCCCAACCATGTGCCCGTGACAAATGGTGGC 3739
Qy 3721 TGCTCCCAATCTGTATTTGCCAAGGGTGATGGGACACCAAGGTGTCTATGCCAGTCAC 3780
Db 3740 TGCTCCCAATCTGTATTTGCCAAGGGTGATGGGACACCAAGGTGTCTATGCCAGTCAC 3799
Qy 3781 CTCTGCTCTCTCGCAAACTCTGTGACCTGTGGAGAGCGCCCACTGTCTCCCGGACAG 3840
Db 3800 CTCTGCTCTCTCGCAAACTCTGTGACCTGTGGAGAGCGCCCACTGTCTCCCGGACAG 3859
Qy 3841 TTGTGATGTGCCACAGGGAGATCGATGTATCTCCCGGGGCTGTGCGTGTGACGGCTTT 3900
Db 3860 TTGTGATGTGCCACAGGGAGATCGATGTATCTCCCGGGGCTGTGCGTGTGACGGCTTT 3919
Qy 3901 CCCGAGTGGATGACACAGAGCGAGAGGCGTCCCGCGTGTGCTCCCGCGCCAGTTC 3960
Db 3920 CCCGAGTGGATGACACAGAGCGAGAGGCGTCCCGCGTGTGCTCCCGCGCCAGTTC 3979
Qy 3961 CCCTGCGCGCGGGTCACTGTGTGACCTGTGCGCTGCGCTGCGAGCGGAGGAGACGT 4020
Db 3980 CCCTGCGCGCGGGTCACTGTGTGACCTGTGCGCTGCGCTGCGAGCGGAGGAGACGT 4039
Qy 4021 CAGGACCGCTCAGAGCGAGCGAGTGTGACCGCATCTGCTGCGCCCAACAGTTCCGGTGT 4080
Db 4040 CAGGACCGCTCAGAGCGAGCGAGTGTGACCGCATCTGCTGCGCCCAACAGTTCCGGTGT 4099
Qy 4081 CGGAGCGGCGAGTGTGCTCATCAACAGAGTGTGAGTCTCTTCCCGCGATGTATCGAC 4140
Db 4100 CGGAGCGGCGAGTGTGCTCATCAACAGAGTGTGAGTCTCTTCCCGCGATGTATCGAC 4159
Qy 4141 GGCTCCGAGCAGTCTCATGTGTGAATACCAAGCGCGCTCAGAGCAGCAGCGCGCCAC 4200
Db 4160 GGCTCCGAGCAGTCTCATGTGTGAATACCAAGCGCGCTCAGAGCAGCAGCGCGCCAC 4219
Qy 4201 AGCAGTGCATCGGCGCGCTCATTTGGCATCATCTCTCTCTCTGCTATGGGTGGTGC 4260
Db 4220 AGCAGTGCATCGGCGCGCTCATTTGGCATCATCTCTCTCTCTGCTATGGGTGGTGC 4279
Qy 4261 TATTTTGTGTGCCAGCGCGTGTGTGCGAGCGTATGCGGGGCGCAACGGGCGCTTCCCG 4320
Db 4280 TATTTTGTGTGCCAGCGCGTGTGTGCGAGCGTATGCGGGGCGCAACGGGCGCTTCCCG 4339
Qy 4321 CACGAGTATGTACGCGGACCCCGCAGTGCCTCAATTTTCATAGCCCGGGCGGTTC 4380
Db 4340 CACGAGTATGTACGCGGACCCCGCAGTGCCTCAATTTTCATAGCCCGGGCGGTTC 4399
Qy 4381 CAGCATGCCCCCTTTCACAGGATCGCATCGCGAAAGTCCATGATGAGTCCCGTGAAGCTG 4440
Db 4400 CAGCATGCCCCCTTTCACAGGATCGCATCGCGAAAGTCCATGATGAGTCCCGTGAAGCTG 4459
Qy 4441 ATGGGGGGCGGGGGGGTGCCTCTGTACGACCGGAACACGTCACAGGGGCGCTCGTTC 4500
Db 4460 ATGGGGGGCGGGGGGGTGCCTCTGTACGACCGGAACACGTCACAGGGGCGCTCGTTC 4519
Qy 4501 AGCAGTCTGTCCAGCAGGAGCGCAGCTGTACCGCGGATCTGAAACCGCGCGCTCC 4560
Db 4520 AGCAGTCTGTCCAGCAGGAGCGCAGCTGTACCGCGGATCTGAAACCGCGCGCTCC 4579
Qy 4561 CCGGCCAGGACCCCTCCCTGTACAAATGACATGTTCTACTCTTCAAAATTCGCGCC 4620
Db 4580 CCGGCCAGGACCCCTCCCTGTACAAATGACATGTTCTACTCTTCAAAATTCGCGCC 4639
Qy 4621 ACTGCGAGACCGTACAGGCCCTATCATCTCAGGAATGGCGCCCGACGACGCCCTTCG 4680
Db 4640 ACTGCGAGACCGTACAGGCCCTATCATCTCAGGAATGGCGCCCGACGACGCCCTTCG 4699

Qy 4681 AGCACCAGCTGTGTGACAGCGACTACAGCCGCGCAGCGCTGMAAGGCCAGCAAGTACTAC 4740
Db 4700 AGCACCAGCTGTGTGACAGCGACTACAGCCGCGCAGCGCTGMAAGGCCAGCAAGTACTAC 4759
Qy 4741 CTGGATTTGAACTCGGACTCGAGACCCCTATCCACCCCGCCACCGAGAGGAGTCTTCCATCTC 4800
Db 4760 CTGGATTTGAACTCGGACTCGAGACCCCTATCCACCCCGCCACCGAGAGGAGTCTTCCATCTC 4819
Qy 4801 CTGTGCGGAGGAGCAGCTGCGCGCTCGCGCGCCACCGAGAGGAGTCTTCCATCTC 4860
Db 4820 CTGTGCGGAGGAGCAGCTGCGCGCTCGCGCGCCACCGAGAGGAGTCTTCCATCTC 4879
Qy 4861 TTCCGCGCCCTCCGCTCCCGCTGACGAGTCTATCTGACCTCGCGCGCCACCTCTGCG 4920
Db 4880 TTCCGCGCCCTCCGCTCCCGCTGACGAGTCTATCTGACCTCGCGCGCCACCTCTGCG 4939
Qy 4921 TTCTCTGTGCGCCCTGTAATAGTTTTTAAATATGAACAAAGAAAAATATATTTTATGAT 4980
Db 4940 TTCTCTGTGCGCCCTGTAATAGTTTTTAAATATGAACAAAGAAAAATATATTTTATGAT 4999
Qy 4981 TTAAAAATAAATAATATTGGATTTTAAAAACATGAGAAATGTGAATCTGTGATGGGTG 5040
Db 5000 TTAAAAATAAATAATATTGGATTTTAAAAACATGAGAAATGTGAATCTGTGATGGGTG 5059
Qy 5041 GGCAGGGCTGGGAGAACTTTTGTGA 5063
Db 5060 GGCAGGGCTGGGAGAACTTTTGTGA 5082

RESULT 6
ACC45358
ID ACC45358 standard; DNA; 5120 BP.
XX ACC45358;
XX 02-JUN-2003 (first entry)
XX Human wild-type LRP5 gene.
XX Human; high bone mass; HBM; LRP5; LRP6; transgenic; bone mass modulation;
KW gene therapy; bone density modulation; bone strength; trabecular number;
KW bone size; bone tissue connectivity; bone disease; osteoporosis;
KW osteomalacia; rickets; Paget's disease; neoplasm of the bone; gene; ds.
XX Homo sapiens.
OS
XX WO200292764-A2.
FN
XX 21-NOV-2002.
PD
XX 13-MAY-2002; 2002WO-US014876.
PF
XX 11-MAY-2001; 2001US-0290071P.
PR 17-MAY-2001; 2001US-0291311P.
PR 01-FEB-2002; 2002US-0353058P.
PR 04-MAR-2002; 2002US-0361293P.
XX (GENO-) GENOME THERAPEUTICS CORP.
PA (AMHP) WYETH.
PA
PI Babij P, Bex FJ, Yaworsky FJ, Bodine PV;
XX
DR WPI; 2003-129278/12.
DR P-PSDB; ABR41093.
XX
PT New transgenic animals (e.g. mice), useful as models for studying bone
PT density modulation, developing drugs for treating or preventing bone
PT diseases (e.g. osteoporosis), or diagnosing diseases characterized by
PT reduced bone density.
PS Disclosure; Fig 6; 603pp; English.
XX

1801 CAGCTGCCGACCTGATGGGGCTCAAAGCTGTGAATGTGGCCAAAGTCTGTGGAAACCAAC 1860
1820 CAGCTGCCGACCTGATGGGGCTCAAAGCTGTGAATGTGGCCAAAGTCTGTGGAAACCAAC 1879
1861 CCGTGTGCGGACAGAAACCGGGGGTGCACGACCTGTGCTTCTTCAACCCCAACGCAACC 1920
1880 CCGTGTGCGGACAGAAACCGGGGGTGCACGACCTGTGCTTCTTCAACCCCAACGCAACC 1939
1921 CGGTGTGGCTGCCCATCGGCTGTGAGCTGTGAGTGATGATGAAGACCTGTGATCGTGCT 1980
1940 CGGTGTGGCTGCCCATCGGCTGTGAGCTGTGAGTGATGATGAAGACCTGTGATCGTGCT 1999
1981 GAGGCTTCTTGGTCTTCAACAGACAGACGCGCATCCACAGGATCTCCCTCGAGACCAAT 2040
2000 GAGGCTTCTTGGTCTTCAACAGACAGACGCGCATCCACAGGATCTCCCTCGAGACCAAT 2059
2041 AACAAACGAGCTGGGCGATCCCGCTCAACGGGCGTCAAGGAGGCTCAGCCCTGGACTTTGAT 2100
2060 AACAAACGAGCTGGGCGATCCCGCTCAACGGGCGTCAAGGAGGCTCAGCCCTGGACTTTGAT 2119
2101 GTGTCCAAACCAACATCTACTGGACAGACGTCAGCCTGAAGACCATCAGCGCGGCTTC 2160
2120 GTGTCCAAACCAACATCTACTGGACAGACGTCAGCCTGAAGACCATCAGCGCGGCTTC 2179
2161 ATGAACGGGAGCTCGTGGAGCAGCTGTGGAGTTTGGCTTGTACTACCCCGAGGGCATG 2220
2180 ATGAACGGGAGCTCGTGGAGCAGCTGTGGAGTTTGGCTTGTACTACCCCGAGGGCATG 2239
2221 GCCGTGTGACTGGATGGGCAAGAACCTTACTTGGGCGGACACTGGGACCAACAGAAATCGAA 2280
2240 GCCGTGTGACTGGATGGGCAAGAACCTTACTTGGGCGGACACTGGGACCAACAGAAATCGAA 2299
2281 GTGGCGGGCTGGACGGGAGTTCCGGCAAGTCTCGTGTGGAGGACTTGGACAAACCCG 2340
2300 GTGGCGGGCTGGACGGGAGTTCCGGCAAGTCTCGTGTGGAGGAGTTGGACAAACCCG 2359
2341 AGGTGCTGGCGCTGGATCCCAACAAAGGGCTACATCTACTGGACCGAGTGGGGGGCAAG 2400
2360 AGGTGCTGGCGCTGGATCCCAACAAAGGGCTACATCTACTGGACCGAGTGGGGGGCAAG 2419
2401 CCGAGGATGTGCGGGCTTTCATGSAACGGGACCAACTGATGATGATGATGATGATGATG 2460
2420 CCGAGGATGTGCGGGCTTTCATGSAACGGGACCAACTGATGATGATGATGATGATGATG 2479
2461 GCGCGGGCAACGACCTCAACATTTGACTACGCTGACGAGCGCTTACTTGAACCGACTG 2520
2480 GCGCGGGCAACGACCTCAACATTTGACTACGCTGACGAGCGCTTACTTGAACCGACTG 2539
2521 GACACCAACATGATCGAGTCTGCAACATGCTGGGTCAAGGAGCGGTCTGATTTGGCGAC 2580
2540 GACACCAACATGATCGAGTCTGCAACATGCTGGGTCAAGGAGCGGTCTGATTTGGCGAC 2599
2581 GATCTCCGCAACCGCTTCCGCTGACGAGTACAGCGATTATATCTACTGGACAGACTGG 2640
2600 GATCTCCGCAACCGCTTCCGCTGACGAGTACAGCGATTATATCTACTGGACAGACTGG 2659
2641 AATCTGACACGANTGACGGGCGGACAAAGACTAGCGGCGGGAACCGACCCCTCATCCAG 2700
2660 AATCTGACACGANTGACGGGCGGACAAAGACTAGCGGCGGGAACCGACCCCTCATCCAG 2719
2701 GGCACCTTGAATCTGATGAGACATCTCTGTGTGTTTCACTCTCCCGCAGGATGGCTTC 2760
2720 GGCACCTTGAATCTGATGAGACATCTCTGTGTGTTTCACTCTCCCGCAGGATGGCTTC 2779
2761 AATGATGTATGCAACAAACGGGCGAGTGTGGGAGCTGTGCTTGGCATCCCGGGCGGC 2820
2780 AATGATGTATGCAACAAACGGGCGAGTGTGGGAGCTGTGCTTGGCATCCCGGGCGGC 2839
2821 CACGCTCGGCTGGCTCTCACTACCTTGAACCCGAGCGGCAATGTGACGCCG 2880
2840 CACGCTCGGCTGGCTCTCACTACCTTGAACCCGAGCGGCAATGTGACGCCG 2899

2881 CCCACCACTTCTTGTCTTTCAGCCAGAAATCTGCCATCAGTCCGATGATCCCGACGAC 2940
2900 CCCACCACTTCTTGTCTTTCAGCCAGAAATCTGCCATCAGTCCGATGATCCCGACGAC 2959
2941 CAGCACAGCCCGGATCTCATCTCTGCTTGCATGAGTGTGAGGAACGTCAAAGCCATCGAC 3000
2960 CAGCACAGCCCGGATCTCATCTCTGCTTGCATGAGTGTGAGGAACGTCAAAGCCATCGAC 3019
3001 TATGACCACTGTGACAGATTTCTACTCTGGTGGATGGGCGGCGAGAAATCAGAGGAGCC 3060
3020 TATGACCACTGTGACAGATTTCTACTCTGGTGGATGGGCGGCGAGAAATCAGAGGAGCC 3079
3061 AAGGACAGCGGAGCCAGCCCTTGTGTTTGTAGCTCTCTGAGCCAGAGCCAAACCCAGAC 3120
3080 AAGGACAGCGGAGCCAGCCCTTGTGTTTGTAGCTCTCTGAGCCAGAGCCAAACCCAGAC 3139
3121 AGGACAGCCCAACGACTCAGCATCTACAGCCGGAACATGTTTCTGACGTGCGAG 3180
3140 AGGACAGCCCAACGACTCAGCATCTACAGCCGGAACATGTTTCTGACGTGCGAG 3199
3181 GCCACCAATACCATCAAGTCCACAGGCTGAGCGGGAGGACATGGGGGTGTGCTGCT 3240
3200 GCCACCAATACCATCAAGTCCACAGGCTGAGCGGGAGGACATGGGGGTGTGCTGCT 3259
3241 GGGGACCGGACAAAGCCAGGGCCATCGTCTCAACGGGAGCGAGGGTACTCTGTACTTC 3300
3260 GGGGACCGGACAAAGCCAGGGCCATCGTCTCAACGGGAGCGAGGGTACTCTGTACTTC 3319
3301 ACCAACTATGACGAGCCGGGACGCAAGATCGAACCGCAGCCCTGGAGCGGACCCAGGCGC 3360
3320 ACCAACTATGACGAGCCGGGACGCAAGATCGAACCGCAGCCCTGGAGCGGACCCAGGCGC 3379
3361 GAGTCTCTTTCACACCGGCTCATCGGCCCTGTGGCCCTGGTGGTGGACAAACACTG 3420
3380 GAGTCTCTTTCACACCGGCTCATCGGCCCTGTGGCCCTGGTGGTGGACAAACACTG 3439
3421 GCGAAGCTGTCTGGGTGGACCGGACCTGAAGCGCATTTAGAGTGTGTGACTGTGACGG 3480
3440 GCGAAGCTGTCTGGGTGGACCGGACCTGAAGCGCATTTAGAGTGTGTGACTGTGACGG 3499
3481 GCCAACCGCTGTACCTTGGAGGACCGCAACATCTGTGACGCTCTGGGCTGTGACCAATCTT 3540
3500 GCCAACCGCTGTACCTTGGAGGACCGCAACATCTGTGACGCTCTGGGCTGTGACCAATCTT 3559
3541 GCGAAGCATCTCTACTGATCGACCGGCGGACGAGATGATGAGCGGTGTGGAGAGACC 3600
3560 GCGAAGCATCTCTACTGATCGACCGGCGGACGAGATGATGAGCGGTGTGGAGAGACC 3619
3601 ACCGGGACAAAGCGGACTCGCATCCAGGGCGGTGTGGCCCACTCCTGCGCATCCATGCA 3660
3620 ACCGGGACAAAGCGGACTCGCATCCAGGGCGGTGTGGCCCACTCCTGCGCATCCATGCA 3679
3661 GTGAGGAAGTACGCTTGGAGGAGTTCTAGCCCAACCATGTGCCCGTGAACAATGGTGGC 3720
3680 GTGAGGAAGTACGCTTGGAGGAGTTCTAGCCCAACCATGTGCCCGTGAACAATGGTGGC 3739
3721 TGCTCCCACTCTGTATTTGCCAAGGGTGTGGGACACCAACGGTGTCTATGCGCATCCAC 3780
3740 TGCTCCCACTCTGTATTTGCCAAGGGTGTGGGACACCAACGGTGTCTATGCGCATCCAC 3799
3781 CTGCTGTCTCTGCAAAACCTGTGACCTGTGGAGAGCGGCGCACTGTCTCCCGGACGAG 3840
3800 CTGCTGTCTCTGCAAAACCTGTGACCTGTGGAGAGCGGCGCACTGTCTCCCGGACGAG 3859
3841 TTTGCAATGTGCAACAGGGGAGATCGACTGTATCCCGGGGCTGTGGCTGTGACGGCTTT 3900
3860 TTTGCAATGTGCAACAGGGGAGATCGACTGTATCCCGGGGCTGTGGCTGTGACGGCTTT 3919
3901 CCGGAGTCCGATGACCAAGAGGACGAGAGGGCTGTGCCCTGTGCTCGCCCGCCAGTTC 3960
3920 CCGGAGTCCGATGACCAAGAGGAGGGGCTGTGCCCTGTGCTCGCCCGCCAGTTC 3979
3961 CCCTGCGCGGGGTCTAGTGTGGACCTGTGCGCTGTGGCTGTGGAGCGGCGAGGAGTGT 4020

Db 200 CGGCTGGTGGACGCGCGGAGTCAAGCTGAGTCCACCATCTGCTGAGTGGAG 259
Qy 241 GATCGGCGCAGTGGACTTCCAGTCTTCAAGGAGCGGTGTAATGGACAGAGTGGC 300
Db 260 GATCGGCGCAGTGGACTTCCAGTCTTCAAGGAGCGGTGTAATGGACAGAGTGGC 319
Qy 301 GAGGAGGCATCAAGCAGACCTACTGAAACAGACGCGGCGCGCTGTCAGAACTGGTC 360
Db 320 GAGGAGGCATCAAGCAGACCTACTGAAACAGACGCGGCGCGCTGTCAGAACTGGTC 379
Qy 361 ATCTCCGCGCTGGTCTCTCCCGAGCGCTCGCTCGAGTGGGTGGGCAAGAGTGTAC 420
Db 380 ATCTCCGCGCTGGTCTCTCCCGAGCGCTCGCTCGAGTGGGTGGGCAAGAGTGTAC 439
Qy 421 TGGACGACTCAGAGACCAACCGCATCGAGTGGCACTCAATGGCACATCCCGGAAG 480
Db 440 TGGACGACTCAGAGACCAACCGCATCGAGTGGCACTCAATGGCACATCCCGGAAG 499
Qy 481 GTGCTCTTGGCAGGACCTTTGACCAGCTTAGGGCCATCGCTTTGGACCCCGCTCACGGG 540
Db 500 GTGCTCTTGGCAGGACCTTTGACCAGCTTAGGGCCATCGCTTTGGACCCCGCTCACGGG 559
Qy 541 TACATGTACTGGACAGACTGGGTGGAGCGCCCGGATTTGAGCGGGCAGGATGGTGGC 600
Db 560 TACATGTACTGGACAGACTGGGTGGAGCGCCCGGATTTGAGCGGGCAGGATGGTGGC 619
Qy 601 AGCACCGGAAGATCATTTGTGGACTCGGACATTTACTGGCCCAATGACTGACCATCGAC 660
Db 620 AGCACCGGAAGATCATTTGTGGACTCGGACATTTACTGGCCCAATGACTGACCATCGAC 679
Qy 661 CTGGAGGAGCAGAACTCTACTGGGCTGACGCCAAGCTCAGCTTTCATCCACCGTGCACAC 720
Db 680 CTGGAGGAGCAGAACTCTACTGGGCTGACGCCAAGCTCAGCTTTCATCCACCGTGCACAC 739
Qy 721 CTGACCGGCTGCTCCGCGCAGAGGTGGTGGAGGGCAGCTGACGACCCCTTGGCCCTG 780
Db 740 CTGACCGGCTGCTCCGCGCAGAGGTGGTGGAGGGCAGCTGACGACCCCTTGGCCCTG 799
Qy 781 ACGCTCTCCGGGACACTCTGTACTGGACAGACTGGCAGACCCGCTCCATCCATGCCCTGC 840
Db 800 ACGCTCTCCGGGACACTCTGTACTGGACAGACTGGCAGACCCGCTCCATCCATGCCCTGC 859
Qy 841 AACAAAGCGCATGGGGGGAAGAGGAGATCTGTAGTGGCCCTCTACTCACCCATGGAC 900
Db 860 AACAAAGCGCATGGGGGGAAGAGGAGATCTGTAGTGGCCCTCTACTCACCCATGGAC 919
Qy 901 ATCCAGTGTGACGACGAGCGCAGCTTTCTTCCACACTCGCTGTGAGGAGGACAAT 960
Db 920 ATCCAGTGTGACGACGAGCGCAGCTTTCTTCCACACTCGCTGTGAGGAGGACAAT 979
Qy 961 GCGGCTGTCCCACTGTGCTGTCTGTCCCAAGCGAGCTTTCTACATGCGCTGC 1020
Db 980 GCGGCTGTCCCACTGTGCTGTCTGTCCCAAGCGAGCTTTCTACATGCGCTGC 1039
Qy 1021 CCCACGGGTGTGACGTGACGAGCAACCGGAGGACGTGTAAAGCAGGAGCGGAGGAGTG 1080
Db 1040 CCCACGGGTGTGACGTGACGAGCAACCGGAGGACGTGTAAAGCAGGAGCGGAGGAGTG 1099
Qy 1081 CTGCTGTGCGCGCGGAGAGGACCTAGGAGGATCTCGCTGGACAGCGCGGACTTCAAC 1140
Db 1100 CTGCTGTGCGCGCGGAGAGGACCTAGGAGGATCTCGCTGGACAGCGCGGACTTCAAC 1159
Qy 1141 GACATCGTGTGAGGTGGACGACATCCGACGCCATTTGCATCGACTACGACCCGCTA 1200
Db 1160 GACATCGTGTGAGGTGGACGACATCCGACGCCATTTGCATCGACTACGACCCGCTA 1219
Qy 1201 GAGGCTATGTCTACTGGACAGATGAGGAGTGGGGCCATCCGACGAGGGGTACTGGAC 1260
Db 1220 GAGGCTATGTCTACTGGACAGATGAGGAGTGGGGCCATCCGACGAGGGGTACTGGAC 1279
Qy 1261 GGGTCTGGGGCGCAGACGTGGTCAACACCGAGATCAACGACCCCGATGGCATCGCGTC 1320
Db 1280 GGGTCTGGGGCGCAGACGTGGTCAACACCGAGATCAACGACCCCGATGGCATCGCGTC 1339

Qy 1321 GACTGGGTGGCCCGAABACCTTACTTGACCGACACGGGCGACCGCATCGAGGTGACG 1380
Db 1340 GACTGGGTGGCCCGAABACCTTACTTGACCGACACGGGCGACCGCATCGAGGTGACG 1399
Qy 1381 CGCTCAACGGCAGCTCCCGCAAGATCTTGTGTGCGAGGACCTTGGAGAGCCCGGAGCC 1440
Db 1400 CGCTCAACGGCAGCTCCCGCAAGATCTTGTGTGCGAGGACCTTGGAGAGCCCGGAGCC 1459
Qy 1441 ATCGCACTGACACCCCGTGTGGGCTCATGTACTGGAAGAGACTGGGAGAGAAACCTTAA 1500
Db 1460 ATCGCACTGACACCCCGTGTGGGCTCATGTACTGGAAGAGACTGGGAGAGAAACCTTAA 1519
Qy 1501 ATCGAGTGTGCAACTTGGATGGGACAGCGGTGTGTGTGTCAATGCGCTCCCTCGG 1560
Db 1520 ATCGAGTGTGCAACTTGGATGGGACAGCGGTGTGTGTGTCAATGCGCTCCCTCGG 1579
Qy 1561 TGGCCCAACCGCTTGGCCCTTGGACCTCGAGAGGGGAGCTTACTTGGGGAGAGCCCAAG 1620
Db 1580 TGGCCCAACCGCTTGGCCCTTGGACCTCGAGAGGGGAGCTTACTTGGGGAGAGCCCAAG 1639
Qy 1621 ACAGACAAGATCGAGGTGATCAATGTTGATGGGACAGAGGCGGACCTTCTTGGAGAC 1680
Db 1640 ACAGACAAGATCGAGGTGATCAATGTTGATGGGACAGAGGCGGAGCCCTCTTGGAGAC 1699
Qy 1681 AAGCTCCGCGACATTTTGGGTTTCACTGCTGGGGGACTTCTACTTCTGAGCTGACTGG 1740
Db 1700 AAGCTCCGCGACATTTTGGGTTTCACTGCTGGGGGACTTCTACTTCTGAGCTGACTGG 1759
Qy 1741 CAGCGCGCAGCATCGAGCGGGTGCAACAAGTCAAGGTCAGGCGGAGCTCATCATTTGAC 1800
Db 1760 CAGCGCGCAGCATCGAGCGGGTGCAACAAGTCAAGGTCAGGCGGAGCTCATCATTTGAC 1819
Qy 1801 CAGTGTGCGGACCTGATGGGGCTCAAGCTGTGTAATGTGGCCAAAGTGTGCGAACCAAC 1860
Db 1820 CAGTGTGCGGACCTGATGGGGCTCAAGCTGTGTAATGTGGCCAAAGTGTGCGAACCAAC 1879
Qy 1861 CCGTGTGCGGACAGGAAACGGGGGTGAGCCACCTGCTGCTTTCACACCCACGCAACC 1920
Db 1880 CCGTGTGCGGACAGGAAACGGGGGTGAGCCACCTGCTGCTTTCACACCCACGCAACC 1939
Qy 1921 CCGTGTGCGCTGCCCCATCGGCTTGGAGCTGTGAGTGACATGAAGACTTGCATCGTGCCT 1980
Db 1940 CCGTGTGCGCTGCCCCATCGGCTTGGAGCTGTGAGTGACATGAAGACTTGCATCGTGCCT 1999
Qy 1981 GAGGCTTCTTGTGCTTTCACAGCAGAGCGCCATCCACAGGATCTCCCTCGAGACCAAT 2040
Db 2000 GAGGCTTCTTGTGCTTTCACAGCAGAGCGCCATCCACAGGATCTCCCTCGAGACCAAT 2059
Qy 2041 AACAAACGCTGGCCATCCCGCTCACGGGCGTCAAGGAGGCTCAGGCTTGGACTTTGAT 2100
Db 2060 AACAAACGCTGGCCATCCCGCTCACGGGCGTCAAGGAGGCTCAGGCTTGGACTTTGAT 2119
Qy 2101 GTGTCCAAACCAATCTACTGGAAGAGCTCAGCTGAGGAGCCATCAGCGCGGCTTC 2160
Db 2120 GTGTCCAAACCAATCTACTGGAAGAGCTCAGCTGAGGAGCCATCAGCGCGGCTTC 2179
Qy 2161 ATGAAACGGGAGCTCGGTGGAGCAGTGTGGAGTTTGGCCCTTGTACTACCCCGAGGAGTG 2220
Db 2180 ATGAAACGGGAGCTCGGTGGAGCAGTGTGGAGTTTGGCCCTTGTACTACCCCGAGGAGTG 2239
Qy 2221 GCGCTTGTGATGGCAAGAACTCTACTTGGGCGGACACTGGGAGCCAAACAGAAATCGAA 2280
Db 2240 GCGCTTGTGATGGCAAGAACTCTACTTGGGCGGACACTGGGAGCCAAACAGAAATCGAA 2299
Qy 2281 GTGCGCGGTGAGCGGCGAGTTCCGCAAGTCTCTGTGTGGAGGAGCTTGGACAAACCG 2340
Db 2300 GTGCGCGGTGAGCGGCGAGTTCCGCAAGTCTCTGTGTGGAGGAGCTTGGACAAACCG 2359
Qy 2341 AGTGTGCTGCGCTTGGATCCCAACAGGGCTCATCTACTTGGACCGGAGTGGGGCGGAG 2400
Db 2360 AGTGTGCTGCGCTTGGATCCCAACAGGGCTCATCTACTTGGACCGGAGTGGGGCGGAG 2419

Db 680 CTGGAGGACGAGAAGCTCTACTGGGCTGACGCCAAGCTCAGCTTCATCCACCGTGCCAAAC 739
QY 721 CTGGACGGCTCGTTCCGGCAGAGGTGTGGAGGCGAGCTGACGCAACCCCTTCGCCCTG 780
Db 740 CTGGACGGCTCGTTCCGGCAGAGGTGTGGAGGCGAGCTTACGACACCCCTTCGCCCTG 799
QY 781 ACGCTCTCGGGGACACTCTGTACTGGACAGACTGGCAGACCCGCTCCATCCATCCCTGC 840
Db 800 ACGCTCTCGGGGACACTCTGTACTGGACAGACTGGCAGACCCGCTCCATCCATCCCTGC 859
QY 841 AACAGCGCACTGGGGGAAAGGAGAGATCTTGAGTGCCTCTACTCAACCCATGGAC 900
Db 860 AACAGCGCACTGGGGGAAAGGAGAGATCTTGAGTGCCTCTACTCAACCCATGGAC 919
QY 901 ATCCAGGTGCTCAGCAGGAGCGGAGCTTTCTTCCACATCTCGCTGTGAGGAGACAAT 960
Db 920 ATCCAGGTGCTCAGCAGGAGCGGAGCTTTCTTCCACATCTCGCTGTGAGGAGACAAT 979
QY 961 GCGGCTGCTCCCACTGTGCTGTCTCCCAAGCGAGCCTTTCTACACATGCGCCTGC 1020
Db 980 GCGGCTGCTCCCACTGTGCTGTCTCCCAAGCGAGCCTTTCTACACATGCGCCTGC 1039
QY 1021 CCCAGCGGTGTGAGCTGCGAGGACAAACCGCAGAGCTGTAAAGCAGGAGCGGAGAGGTG 1080
Db 1040 CCCAGCGGTGTGAGCTGCGAGGACAAACCGCAGAGCTGTAAAGCAGGAGCGGAGAGGTG 1099
QY 1081 CTGCTGCTGGCCCGCGAGCGACCTACGAGGATCTCGCTGGACACCGCGGACTTCACC 1140
Db 1100 CTGCTGCTGGCCCGCGAGCGACCTACGAGGATCTCGCTGGACACCGCGGACTTCACC 1159
QY 1141 GACATCGTGTGAGGTGGAAGACATCCGGCAGCGCATTTGCCATCGACTAGCACCCGCTA 1200
Db 1160 GACATCGTGTGAGGTGGAAGACATCCGGCAGCGCATTTGCCATCGACTAGCACCCGCTA 1219
QY 1201 GAGGCTATGTCTACTGACAGATGACGAGTGGGGCCATCCGACAGCGGTACTGGAC 1260
Db 1220 GAGGCTATGTCTACTGACAGATGACGAGTGGGGCCATCCGACAGCGGTACTGGAC 1279
QY 1261 GGGTCTGGGGCGCAGACCTGGTCAACACGAGATCAACGACCCCGATGGCATCGCGTC 1320
Db 1280 GGGTCTGGGGCGCAGACCTGGTCAACACGAGATCAACGACCCCGATGGCATCGCGTC 1339
QY 1321 GACTGGGTGGCCGAAACCTCTACTGGACCGACACCGGCAACCGCATCGAGTGAAG 1380
Db 1340 GACTGGGTGGCCGAAACCTCTACTGGACCGACACCGGCAACCGCATCGAGTGAAG 1399
QY 1381 CGCTCAACGGGCACTCCCGCAAGATCTCTGTGTGCGAGGACCTTGAACGAGCGCCGAGCC 1440
Db 1400 CGCTCAACGGGCACTCCCGCAAGATCTCTGTGTGCGAGGACCTTGAACGAGCGCCGAGCC 1459
QY 1441 ATCGCACTGCAACCCGTCATGGGCTCATGTACTGGACAGACTGGGAGAGAACCTTAA 1500
Db 1460 ATCGCACTGCAACCCGTCATGGGCTCATGTACTGGACAGACTGGGAGAGAACCTTAA 1519
QY 1501 ATCGAGTGTGCAACTTGGATGGGAGGAGCGGCTGTGCTGCTCAATGCCCTCCCTCGGG 1560
Db 1520 ATCGAGTGTGCAACTTGGATGGGAGGAGCGGCTGTGCTGCTCAATGCCCTCCCTCGGG 1579
QY 1561 TGGCCCAACGGGCTGGCCCTGCACTGTGAGAGGAGAGCTCTACTGGGAGAGAGCCCAAG 1620
Db 1580 TGGCCCAACGGGCTGGCCCTGCACTGTGAGAGGAGAGCTCTACTGGGAGAGAGCCCAAG 1639
QY 1621 ACAGACAAGATCGAGTGCATCAATGTTGATGGGAGAGAGCGGACCCCTCTCGAGGAC 1680
Db 1640 ACAGACAAGATCGAGTGCATCAATGTTGATGGGAGAGAGCGGACCCCTCTCGAGGAC 1699
QY 1681 AAGCTCCCGCACTTTTCGGGTTCAAGCTGTGCGGGGACTTCACTACTTGGAGTGAAGTGG 1740
Db 1700 AAGCTCCCGCACTTTTCGGGTTCAAGCTGTGCGGGGACTTCACTACTTGGAGTGAAGTGG 1759
QY 1741 CAGCGCCGAGATCGAGCGGTGCAAAAGTCAAGGCGAGCGGAGCTCATATTGAC 1800
Db 1760 CAGCGCCGAGATCGAGCGGTGCAAAAGTCAAGGCGAGCGGAGCTCATATTGAC 1819

QY 1801 CAGCTGCCGACCTGATGGGGCTCAAGCTCTGATGTGGCCNAGGTCTGTGGAGACCAAC 1860
Db 1820 CAGTGCCTCCGACCTGATGGGGCTCAAGCTCTGATGTGGCCNAGGTCTGTGGAGACCAAC 1879
QY 1861 CCGTGTGCGGACAGGAAACGGGGGTGAGCACCCTGTGCTTTCTTACACCCACCAACCAAC 1920
Db 1880 CCGTGTGCGGACAGGAAACGGGGGTGAGCACCCTGTGCTTTCTTACACCCACCAACCAAC 1939
QY 1921 CCGTGTGCTGCCCATCGGCTCGAGCTGTGAGTGTGATGATGAAGAACTGTGATCTGCT 1980
Db 1940 CCGTGTGCTGCCCATCGGCTCGAGCTGTGAGTGTGATGATGAAGAACTGTGATCTGCT 1999
QY 1981 GAGGCTTTCTTGGTCTTCCACGAGAGCGGCGCATCCACAGATCTCCCTCGAGACCAAT 2040
Db 2000 GAGGCTTTCTTGGTCTTCCACGAGAGCGGCGCATCCACAGATCTCCCTCGAGACCAAT 2059
QY 2041 AACAAACGAGTGGGCTATCCCGCTCACGGGCGCTCAAGGAGGCTCAGCCCTGAGCTTTGAT 2100
Db 2060 AACAAACGAGTGGGCTATCCCGCTCACGGGCGCTCAAGGAGGCTCAGCCCTGAGCTTTGAT 2119
QY 2101 GTGTCCAAACCAACATCTACTGGACAGACTGTGAGCTGTGAAGAACTGTGAGCGGCTTC 2160
Db 2120 GTGTCCAAACCAACATCTACTGGACAGACTGTGAGCTGTGAAGAACTGTGAGCGGCTTC 2179
QY 2161 ATGAACGGGAGCTCGGTGGAGCAGTGTGTGAGTTTGGCCCTTGACTACCCGAGGGCATG 2220
Db 2180 ATGAACGGGAGCTCGGTGGAGCAGTGTGTGAGTTTGGCCCTTGACTACCCGAGGGCATG 2239
QY 2221 GCCGTGTGACTGGATGGGCAAGAACTCTACTTGGGCGGACACTGGGACCAACAGAACTGAA 2280
Db 2240 GCCGTGTGACTGGATGGGCAAGAACTCTACTTGGGCGGACACTGGGACCAACAGAACTGAA 2299
QY 2281 GTGGCGGCTGTGAGCAGGCGAGTTCCGGCAAGTCTCTGTGTGAGAGGAACTTGTGACAAACCG 2340
Db 2300 GTGGCGGCTGTGAGCAGGCGAGTTCCGGCAAGTCTCTGTGTGAGAGGAACTTGTGACAAACCG 2359
QY 2341 AGGTCTGTGGCCCTGGATCCCAACAGGCTTACATCTGAGCCGAGTGTGGGGGCAAG 2400
Db 2360 AGGTCTGTGGCCCTGGATCCCAACAGGCTTACATCTGAGCCGAGTGTGGGGGCAAG 2419
QY 2401 CCGAGGATCTGTGCGGGCTTTCATGAGCGGGAACCAACTGCATGACGCTGTGTGAGCAAGGTG 2460
Db 2420 CCGAGGATCTGTGCGGGCTTTCATGAGCGGGAACCAACTGCATGACGCTGTGTGAGCAAGGTG 2479
QY 2461 GCGCGGGCAACGACCTTCAACATTTGACTACGCTGACGAGCGGCTTCTACTGAGCGGACCTG 2520
Db 2480 GCGCGGGCAACGACCTTCAACATTTGACTACGCTGACGAGCGGCTTCTACTGAGCGGACCTG 2539
QY 2521 GACACCAACATGATCGAGTCTGCAACATGTGTGGTCAAGGAGCGGCTGTGATTTGCCGAC 2580
Db 2540 GACACCAACATGATCGAGTCTGCAACATGTGTGGTCAAGGAGCGGCTGTGATTTGCCGAC 2599
QY 2581 GATCTCCCGCAACCCGTTGCTGACGAGTACAGCGATTATATCTACTGAGCAAGTGG 2640
Db 2600 GATCTCCCGCAACCCGTTGCTGACGAGTACAGCGATTATATCTACTGAGCAAGTGG 2659
QY 2641 AATCTGACAGCATTTGAGCGGCGGCAAGAGTGTGAGCGGCGGCAACCGCACTTATCCAG 2700
Db 2660 AATCTGACAGCATTTGAGCGGCGGCAAGAGTGTGAGCGGCGGCAACCGCACTTATCCAG 2719
QY 2701 GGGCACCCTGAGCTTCTGTGATGGACATCTGTGTGTTCCACTCTCCCGCAGGATGGCCTC 2760
Db 2720 GGGCACCCTGAGCTTCTGTGATGGACATCTGTGTGTTCCACTCTCCCGCAGGATGGCCTC 2779
QY 2761 AATGACTGTATGCAACAAACGGGCAAGTGTGGGAGCTGTGCTTGCATCCCGCGCGGC 2820
Db 2780 AATGACTGTATGCAACAAACGGGCAAGTGTGGGAGCTGTGCTTGCATCCCGCGCGGC 2839
QY 2821 CACCGCTGGGCTGGCTCAGCTACACCTGTGAGCCCGCAGCAGCGCACTGTGAGCCG 2880
Db 2840 CACCGCTGGGCTGGCTCAGCTACACCTGTGAGCCCGCAGCAGCGCACTGTGAGCCG 2899

Qy	2881	CCACACACCTTCTTGCTGTTTCAGCCAGAAATCTGCCATCAGTCCGATGATCCCGAGCAC	2944
Db	2900	CCCACCACTTCTTGCTGTTTCAGCCAGAAATCTGCCATCAGTCCGATGATCCCGAGCAC	2959
Qy	2941	CAGCACAGCCGGATCTCATCTCTGCCCTGCATCGACTGAGGAACGTCAAAGCCATCGAC	3000
Db	2960	CAGCACAGCCGGATCTCATCTCTGCCCTGCATCGACTGAGGAACGTCAAAGCCATCGAC	3019
Qy	3001	TATGACCCACTGGCAAGTGTTCATCTACTTGGTGGATGGGCCAGAAACATCAAGCGAGCC	3060
Db	3020	TATGACCCACTGGCAAGTGTTCATCTACTTGGTGGATGGGCCAGAAACATCAAGCGAGCC	3079
Qy	3061	AAGGACAGCGGACCCAGCCCTTGTGTTTGTACCTCTCTGAGCCAAAGCCAAAACCCAGAC	3120
Db	3080	AAGGACAGCGGACCCAGCCCTTGTGTTTGTACCTCTCTGAGCCAAAGCCAAAACCCAGAC	3139
Qy	3121	AGGAGGCCCAAGCACTCAGCATCGACATCTACAGCCGGACACTGTGTTCTGAGCGTGGAG	3180
Db	3140	AGGAGGCCCAAGCACTCAGCATCGACATCTACAGCCGGACACTGTGTTCTGAGCGTGGAG	3199
Qy	3181	GCCACCAATACCACTCAACGCTCCAGAGCTGAGCGGGGAAGCCATGGGGGTGGTCTGCGT	3240
Db	3200	GCCACCAATACCACTCAACGCTCCAGAGCTGAGCGGGGAAGCCATGGGGGTGGTCTGCGT	3259
Qy	3241	GGGAGCCGCGACAAAGCCAGGCCATCGTGTCTAAAGCGGAGCGAGGGTACTGTGTTCTTC	3300
Db	3260	GGGAGCCGCGACAAAGCCAGGCCATCGTGTCTAAAGCGGAGCGAGGGTACTGTGTTCTTC	3319
Qy	3301	ACCAACATGAGGACCGGGCAGCAAGATCGAACGCGCAGCCCTGGACGACCCAGCGAGCGC	3360
Db	3320	ACCAACATGAGGACCGGGCAGCAAGATCGAACGCGCAGCCCTGGACGACCCAGCGAGCGC	3379
Qy	3361	GAGTCTCTTTCACCAACCGGCTCATCCGCCCTGTGGCCCTGTGGTGGGACCAACACTG	3420
Db	3380	GAGTCTCTTTCACCAACCGGCTCATCCGCCCTGTGGCCCTGTGGTGGGACCAACACTG	3439
Qy	3421	GGCAAGCTGTTCTGGGTGGACGGGAACTGAAAGCGCAATTGAGAGCTGTGACTGTGAGG	3480
Db	3440	GGCAAGCTGTTCTGGGTGGACGGGAACTGAAAGCGCAATTGAGAGCTGTGACTGTGAGG	3499
Qy	3481	GCCAAACGCTGACCTTGGAGGACGCAACATCGTGAGCCTCTGGGCTGAGCAATCCCTT	3540
Db	3500	GCCAAACGCTGACCTTGGAGGACGCAACATCGTGAGCCTCTGGGCTGAGCAATCCCTT	3559
Qy	3541	GGCAAGCATCTCTACTGGATCGAACCGCAGCAGATGATCGAGCGTGTGGAGAGAGCC	3600
Db	3560	GGCAAGCATCTCTACTGGATCGAACCGCAGCAGATGATCGAGCGTGTGGAGAGAGCC	3619
Qy	3601	ACGGGAGCAAGCGGACTCGGATCAAGGCGCTGTGCCCCCACTCACTGGGATCCATGCA	3660
Db	3620	ACGGGAGCAAGCGGACTCGGATCAAGGCGCTGTGCCCCCACTCACTGGGATCCATGCA	3679
Qy	3661	GTGGAGAACTGAGCTTGGAGGAGTCTCAGCCCCACCATGTGCCGTGACATGTGGC	3720
Db	3680	GTGGAGAACTGAGCTTGGAGGAGTCTCAGCCCCACCATGTGCCGTGACATGTGGC	3739
Qy	3721	TGCTCCCAACATCTGTATTGCGAAGGTGATGGGACACACAGGTGCTCATGCCAGTCCAC	3780
Db	3740	TGCTCCCAACATCTGTATTGCGAAGGTGATGGGACACACAGGTGCTCATGCCAGTCCAC	3799
Qy	3781	CTCGTGTCTCTGCAAACTGCTGACCTGTGGAGAGCGGCCAACCTGTCTCCCGGACCCAG	3840
Db	3800	CTCGTGTCTCTGCAAACTGCTGACCTGTGGAGAGCGGCCAACCTGTCTCCCGGACCCAG	3859
Qy	3841	TTTGTGATGTGCCACAGGGGAGATCGACTGTATCCCCGGGGCCTGGCGCTGTGAGCGCTT	3900
Db	3860	TTTGTGATGTGCCACAGGGGAGATCGACTGTATCCCCGGGGCCTGGCGCTGTGAGCGCTT	3919
Qy	3901	CCGAGTGGCATGACACAGAGCGAGAGGGCTGCCCGTGTGCTCCGCGGCCAGTTC	3960
Db	3920	CCGAGTGGCATGACACAGAGCGAGAGGGCTGCCCGTGTGCTCCGCGGCCAGTTC	3979
Qy	3961	CCCTGCGCGGGGTCAAGTGTGTGGACCTCTCGCTGCTGCGACGCGCAGGACAGTCTG	4020

Db	3980		CCCTGCGCGGGGTCAAGTGTGTGGACCTTCGCGCTGCGCTGCACGCGCGAGGCAGACTGT	4039
Qy	4021		CAGGACCGCTCAGACGAGCGGAGCTGTGACGCCCATCTGCTCGCTGCGCCAAACAGTTCGGGTGT	4080
Db	4040		CAGACCGCTCAGACGAGAGGTGGACTGTGACGCCCATCTGCTCGCTGCGCCAAACAGTTCGGGTGT	4099
Qy	4081		GCAGCGGCGCAGTGTGTCTCATCAAAACAGAGTGGAGTCTCTTCGCCGACTGTATCGAC	4140
Db	4100		GCAGCGGCGCAGTGTGTCTCATCAAAACAGAGTGGAGTCTCTTCGCCGACTGTATCGAC	4159
Qy	4141		GGCTCCGACGAGCTCATGTGTGAAATCACAAGCGCGCCTCAGACGACGCGCGGCCAC	4200
Db	4160		GGCTCCGACGAGCTCATGTGTGAAATCACAAGCGCGCCTCAGACGACGCGCGGCCAC	4219
Qy	4201		AGCAGTGCCATCGGCGCCGCTCATTTGGCATCATCTCTCTCTCTTCGTATGGGTGTGTCT	4260
Db	4220		AGCAGTGCCATCGGCGCCGCTCATTTGGCATCATCTCTCTCTCTTCGTATGGGTGTGTCT	4279
Qy	4261		TATTTTGTGTGCCAGCGCGTGTGTGTGACGCTATGCGGGGCCAAACGGGCGCTTCGCCG	4320
Db	4280		TATTTTGTGTGCCAGCGCGTGTGTGTGACGCTATGCGGGGCCAAACGGGCGCTTCGCCG	4339
Qy	4321		CACAGTATGTACGCGGGAACCGCGCAGTGCCTCTCAATTTATAGCCCCGCGGGGTTC	4380
Db	4340		CACAGTATGTACGCGGGAACCGCGCAGTGCCTCTCAATTTATAGCCCCGCGGGGTTC	4399
Qy	4381		CAGCATGGCCCTTTCACAGGCATCGCATGCGGAAAGTCCATGATGAGTCCGTGAGGCTG	4440
Db	4400		CAGCATGGCCCTTTCACAGGCATCGCATGCGGAAAGTCCATGATGAGTCCGTGAGGCTG	4459
Qy	4441		ATGGGGGGCGGGCGGGGTGCCCTGTACGACCGGAACCAACGCTCAAGGGGCTCGTCC	4500
Db	4460		ATGGGGGGCGGGCGGGGTGCCCTGTACGACCGGAACCAACGCTCAAGGGGCTCGTCC	4519
Qy	4501		AGCAGCTGTTCAGCACGAAGGCCACGCTGTACCGCCGATCTCTGAACCCGCGCCCTCC	4560
Db	4520		AGCAGCTGTTCAGCACGAAGGCCACGCTGTACCGCCGATCTCTGAACCCGCGCCCTCC	4579
Qy	4561		CCGGCCAGGACCCCTCCCTGTCAACATGGACAATGTTCTACTTCTTCAAAACATTCGGGCC	4620
Db	4580		CCGGCCAGGACCCCTCCCTGTCAACATGGACAATGTTCTACTTCTTCAAAACATTCGGGCC	4639
Qy	4621		ACTGCGAGCCGTACAGGCCCTATCATCATTGAGGAATGCGGCCCGCCGACGACGCCCTGC	4680
Db	4640		ACTGCGAGCCGTACAGGCCCTATCATCATTGAGGAATGCGGCCCGCCGACGACGCCCTGC	4699
Qy	4681		AGCACCGAGTGTGACAGCGACTACAGCGCCAGCGCGCTGGGAAGGCCAGCAAGTACTAC	4740
Db	4700		AGCACCGAGTGTGACAGCGACTACAGCGCCAGCGCGCTGGGAAGGCCAGCAAGTACTAC	4759
Qy	4741		CTGGATTTGAACTCGGACTCAGACCCCTATCTCACCCCCACCCACGCCCCACAGCAGTAC	4800
Db	4760		CTGGATTTGAACTCGGACTCAGACCCCTATCTCACCCCCACCCACGCCCCACAGCAGTAC	4819
Qy	4801		CTGTGCGCGGAGGACAGCTGCGCGCCCTCGCCGCCACCGAGAGGAGTACTTCCATCTC	4860
Db	4820		CTGTGCGCGGAGGACAGCTGCGCGCCCTCGCCGCCACCGAGAGGAGTACTTCCATCTC	4879
Qy	4861		TTCCCGCCCTCGTCCCTCGGACGGAATCATCTGTGACCTCGCGCGGGGCACTCTGGC	4920
Db	4880		TTCCCGCCCTCGTCCCTCGGACGGAATCATCTGTGACCTCGCGCGGGGCACTCTGGC	4939
Qy	4921		TTCTCTGTGCGCCCTGTAAATAGTTTTAAATATCAAAAGAAAAAATATATTTTATGAT	4980
Db	4940		TTCTCTGTGCGCCCTGTAAATAGTTTTAAATATATGAAACAAAGAAAAAATATATTTTATGAT	4999
Qy	4981		TTAAAAATAAATATATTTGGGATTTTAAAAACATGAGAAATGTGAATCTGTGATGGGGTG	5040
Db	5000		TTAAAAATAAATATATTTGGGATTTTAAAAACATGAGAAATGTGAATCTGTGATGGGGTG	5059
Qy	5041		GGCAGGCGTGGGAGAACTTTGTA	5063

Db 5060 GGCAGGCTGGAGAACTTTGTA 5082

RESULT 9

ADR16919

ID ADRI6919 standard; DNA; 5120 BP.

AC ADR16919;

DT 04-NOV-2004 (first entry)

DE Human high bone mass gene, wild type allele Zmax1 DNA.

Human; high bone mass; Zmax1; ds; gene; HBM; osteoporosis; chromosome 11q13.3; osteopathic; LDL receptor; bone development; metabolic bone disease; SNP; single nucleotide polymorphism.

XX
 OS Homo sapiens.

XX	Key	Location/Qualifiers
FH		

FT	CDS	71. .4918
----	-----	-----------

```
FT      /*tag= a
```

FT
/product= "Zmax1 protein"

```
FT /transl_except= (pos:530. .532,aa:Lys)
```

```

F1 /note= "Applies to SEQ ID 3 only"
F2 /note= "Applies to SEQ ID 3 only"
F3 /note= "Applies to SEQ ID 3 only"

```

```
FI /transl_except= (pos:986: :988,aa:Trp)
FT /note= "Analyzed for SEC ID 3 only"
```

```

FT  /note= 'Applies to SEQ ID 3 only'
FT  /transl_except= (pos:1055..1057.

```

```
/name= "Applies to SEQ ID 3 only"
```

```
FT /transl_except= (pos:2162. .2164,
```

FT /note= "Applies to SEQ ID 3 only"

FT	allele	replace(582,T)
TT		

```

BT
/* tag= b
/

```

FT
XX
/Standard-

PN US6780609-B1.

PN US6780609-B1.

PD 24-AUG-2004.

XX
PF 05-APR-2000; 2000US-00543771.

PR 13-JAN-1998; 98US-0071449P.

PR 23-OCT-1998; 98US-0105511P.

PR 13-JAN-1999; 99US-00229319.

XX
PA (GENO-) GENOME THERAPEUTICS CORP.

PI Carulli JP, Little RD, Recker RR, Johnson ML;

DR WPI; 2004-623529/60.

DR P-PSDB; ADR16921, ADR17560.

XX New high bone mass gene of chromosome 1.1013.3, encoding protein useful
PT for treating, diagnosing, preventing, or screening for normal and
PT abnormal conditions of bone, including metabolic bone diseases, e.g.
PT osteoporosis.

PS Disclosure; SEQ ID NO 1; 284pp; English.

The invention relates to an isolated amino acid protein sequence selected from an amino acid sequence appearing as ADR16922 or an amino acid sequence comprising or consisting of the extracellular domain of ADR16922(amino acids 23-1385). ADR16922 is encoded by the HBM (high bone mass) allele of the human Xmai1 gene and has sequence similarity to LDL receptors. Also disclosed are nucleic acids, proteins, cloning vectors, expression vectors, transformed hosts, methods of developing pharmaceutical compositions, methods of identifying molecules involved in bone development, and methods of diagnosing and treating diseases involved in bone development. Specifically disclosed is the Xmai1 gene and the high bone mass (HBM) allele on chromosome 11q13.3 encoding ADR16922. The protein is useful for treating, diagnosing, preventing, or screening for normal and abnormal conditions of bone, including metabolic bone diseases, e.g. osteoporosis. The present sequence is the wild type

Qy 961 GCGGCTGCTCCACCTGTGCTGTCCCAAGCGAGCCTTTCTACACATGCGCCTGC 1020
Db 980 GCGGCTGCTCCACCTGTGCTGTCCCAAGCGAGCCTTTCTACACATGCGCCTGC 1039
Qy 1021 CCCAGGGTGTGAGCTGCAGGACAAACGGCAGAGCTGTAAAGGACGAGCCGAGAGGTG 1080
Db 1040 CCCAGGGTGTGAGCTGCAGGACAAACGGCAGAGCTGTAAAGGACGAGCCGAGAGGTG 1099
Qy 1081 CTGCTGTGCGCCGCGGACGAGCCTACGGAAGGATCTCGCTGGACACGCCGAGCTTCACC 1140
Db 1100 CTGCTGTGCGCCGCGGACGAGCCTACGGAAGGATCTCGCTGGACACGCCGAGCTTCACC 1159
Qy 1141 GACATGTGCTGCGAGGTGGACGACATCCGGCAGCCATTGGCCATCGACTACGAGCCGCTA 1200
Db 1160 GACATGTGCTGCGAGGTGGACGACATCCGGCAGCCATTGGCCATCGACTACGAGCCGCTA 1219
Qy 1201 GAGGGCTATGTCTACTGACAGATGACGAGGTGGCGGCCATCCGAGGGCGTTACTGGAC 1260
Db 1220 GAGGGCTATGTCTACTGACAGATGACGAGGTGGCGGCCATCCGAGGGCGTTACTGGAC 1279
Qy 1261 GGGTCTGGGGCGAGACGCTGCTCAACACCGAGATCAACGACCCCGATGGCATCGCGGTC 1320
Db 1280 GGGTCTGGGGCGAGACGCTGCTCAACACCGAGATCAACGACCCCGATGGCATCGCGGTC 1339
Qy 1321 GACTGGGTGGCCCGAAACCTCTACTGGACCGGACCGGACCGGACCGCATCGAGGTGACG 1380
Db 1340 GACTGGGTGGCCCGAAACCTCTACTGGACCGGACCGGACCGGACCGCATCGAGGTGACG 1399
Qy 1381 GCGCTCAACGGGACCTCCGCAAGATCTGTGTGCGAGGACCTTGGAGAGCCCGGAGCC 1440
Db 1400 GCGCTCAACGGGACCTCCGCAAGATCTGTGTGCGAGGACCTTGGAGAGCCCGGAGCC 1459
Qy 1441 ATCGCACTGCAACCGGTGATGGGCTCATGTACTGGACAGACTGGGAGAGAACCTTAAA 1500
Db 1460 ATCGCACTGCAACCGGTGATGGGCTCATGTACTGGACAGACTGGGAGAGAACCTTAAA 1519
Qy 1501 ATCGAGTGTGCAACTTTGGATGGGCGAGGCGGGTGTGTGGTCAATGCTCCCTCGGG 1560
Db 1520 ATCGAGTGTGCAACTTTGGATGGGCGAGGACGGCGTGTGTGGTCAATGCTCCCTCGGG 1579
Qy 1561 TGGCCCAACGGCCTGGCCCTGGACCTTGACAGAGGGGAGGCTCTACTGGGGAGAGCCCAAG 1620
Db 1580 TGGCCCAACGGCCTGGCCCTGGACCTTGACAGAGGGGAGGCTCTACTGGGGAGAGCCCAAG 1639
Qy 1621 ACAGCAAGATCGAGGTGATCAATGTTGATCGGACGAGAGGCGGACCCCTCTCGAGGAC 1680
Db 1640 ACAGCAAGATCGAGGTGATCAATGTTGATGGGACGAGAGGCGGACCCCTCTCGAGGAC 1699
Qy 1681 AAGCTCCCGCACATTTTCGGGTTTCACTGCTGGGGGACTTTCATCTACTGGACTGACTGG 1740
Db 1700 AAGCTCCCGCACATTTTCGGGTTTCACTGCTGGGGGACTTTCATCTACTGGACTGACTGG 1759
Qy 1741 CAGCGCGCAGCATCGAGCGGGTGGACAGGTCAAGGCGAGCGGGAGCTCATCATTTGAC 1800
Db 1760 CAGCGCGCAGCATCGAGCGGGTGGACAGGTCAAGGCGAGCGGGAGCTCATCATTTGAC 1819
Qy 1801 CAGCTGCCCGACCTGATGGGCTCAAAGCTGTGATGTGGCCAAAGTGTGGGACCAAC 1860
Db 1820 CAGCTGCCCGACCTGATGGGCTCAAAGCTGTGATGTGGCCAAAGTGTGGGACCAAC 1879
Qy 1861 CCGTGTGCGGACAGGAACGGGGGTGGACGACCTGTGCTTCTTTCACACCCCAAGCAACC 1920
Db 1880 CCGTGTGCGGACAGGAACGGGGGTGGACGACCTGTGCTTCTTTCACACCCCAAGCAACC 1939
Qy 1921 CCGTGTGCGCTCCCATCGGCTCGAGCTGTGATGTGATGACATGAAGACCTGTGATCGGCT 1980
Db 1940 CCGTGTGCGCTCCCATCGGCTCGAGCTGTGATGTGATGACATGAAGACCTGTGATCGGCT 1999
Qy 1981 GAGGCTTCTTGTGTTTACAGAGAGCGGCGCATCCACAGGATCTCCCTCGAGACCAAT 2040
Db 2000 GAGGCTTCTTGTGTTTACAGAGAGCGGCGCATCCACAGGATCTCCCTCGAGACCAAT 2059
Qy 2041 AACCAACGACGTGGCCATCCCGCTCAACGGGCGCTCAAGGAGGCGCTCAGGCGCTTGTGAT 2100

Db 2060 AACCAACGACGTGGCCATCCCGCTCAAGGAGGCGCTCAGCCCTGACTTTGAT 2119
Qy 2101 GTGTCCAAACAAACACATCTACTGACAGAGCTGAGCCCTGAAGACATCAGCGCGGCTTC 2160
Db 2120 GTGTCCAAACAAACACATCTACTGACAGAGCTGAGCCCTGAAGACATCAGCGCGGCTTC 2179
Qy 2161 ATGAACGGGAGCTCGGTGGAGCAGTGTGAGATTTGGCCCTTGACTACCCCGAGGGCATG 2220
Db 2180 ATGAACGGGAGCTCGGTGGAGCAGTGTGAGATTTGGCCCTTGACTACCCCGAGGGCATG 2239
Qy 2221 GCCGTTGACTGGATGGGCAAGAACCTCTACTGGGCGGACACTGGGACCAACAGATCGAA 2280
Db 2240 GCCGTTGACTGGATGGGCAAGAACCTCTACTGGGCGGACACTGGGACCAACAGATCGAA 2299
Qy 2281 GTGGCGGGCTGGAGCGGGCAGTTCGGCAAGTCTCTGTGTGAGGGACTTGGACAACCG 2340
Db 2300 GTGGCGGGCTGGAGCGGGCAGTTCGGCAAGTCTCTGTGTGAGGGACTTGGACAACCG 2359
Qy 2341 AGTTCGCTGGCCCTGGATCCCAACAGGCTTACTACTGGACCGAGTGGGGCGGCAAG 2400
Db 2360 AGTTCGCTGGCCCTGGATCCCAACAGGCTTACTACTGGACCGAGTGGGGCGGCAAG 2419
Qy 2401 CCGAGGATCGTGGCGGCTTTCATGGACGGGACCAACTGCATGACGCTGGTGACAAAGTG 2460
Db 2420 CCGAGGATCGTGGCGGCTTTCATGGACGGGACCAACTGCATGACGCTGGTGACAAAGTG 2479
Qy 2461 GCGCGGGCAACGACCTTCACTTGGACTACGCTGACCAAGCCCTTACTTGGACCGACCTG 2520
Db 2480 GCGCGGGCAACGACCTTCACTTGGACTACGCTGACCAAGCCCTTACTTGGACCGACCTG 2539
Qy 2521 GACACCAACATGATCGAGTGTGTCACCAATGCTGGTTCAGGAGCGGCTGTGATTTGCCGAC 2580
Db 2540 GACACCAACATGATCGAGTGTGTCACCAATGCTGGTTCAGGAGCGGCTGTGATTTGCCGAC 2599
Qy 2581 GATCTCCCGCACCCGTTTGGTCTGACGAGTACAGCGATTAATATCTACTTGGACAGACTGG 2640
Db 2600 GATCTCCCGCACCCGTTTGGTCTGACGAGTACAGCGATTAATATCTACTTGGACAGACTGG 2659
Qy 2641 AATCTGACAGCATTGAGCGGGCGGACAGACTAGCGCGCGGAAACCGACCTCATCCAG 2700
Db 2660 AATCTGACAGCATTGAGCGGGCGGACAGACTAGCGCGCGGAAACCGACCTCATCCAG 2719
Qy 2701 GGCACCTGACTTCTGTGATGGACATCTGTGTTCACCTCTCCCGCAGAGTGGCCTC 2760
Db 2720 GGCACCTGACTTCTGTGATGGACATCTGTGTTCACCTCTCCCGCAGAGTGGCCTC 2779
Qy 2761 AATGACTGTATGCAAAACAGGGCAGTGTGGCAGCTGTGCTTGGCCATCCCGGCGGC 2820
Db 2780 AATGACTGTATGCAAAACAGGGCAGTGTGGGCGAGCTGTGCTTGGCCATCCCGGCGGC 2839
Qy 2821 CACGCTGCGGCTGGCCTCACACTACACCTGGACCCCGAGCAGCGGCACTGACGCCG 2880
Db 2840 CACGCTGCGGCTGGCCTCACACTACACCTGGACCCCGAGCAGCGGCACTGACGCCG 2899
Qy 2881 CCCACCACTTCTGTGCTTTCAGCCAGAAATCTGCCATCAGTCCGATGATCCCGAGCAG 2940
Db 2900 CCCACCACTTCTGTGCTTTCAGCCAGAAATCTGCCATCAGTCCGATGATCCCGAGCAG 2959
Qy 2941 CAGCACAGCCGGATCTCATCTCTGCCCTTGCATGGAGTGAAGAACTGTCGATGATCCG 3000
Db 2960 CAGCACAGCCGGATCTCATCTCTGCCCTTGCATGGAGTGAAGAACTGTCGATGATCCG 3019
Qy 3001 TATGACCCACTTGGACAAAGTTTCTACTTGGTGGATGGGGCCGAGAACATCAAGGAGCC 3060
Db 3020 TATGACCCACTTGGACAAAGTTTCTACTTGGTGGATGGGGCCGAGAACATCAAGGAGCC 3079
Qy 3061 AAGGACGAGCGGACCCGAGCCCTTGTGCTTGTGACCTCTCTGAGCCAAAGGCAAAACCCAGAC 3120
Db 3080 AAGGACGAGCGGACCCGAGCCCTTGTGCTTGTGACCTCTCTGAGCCAAAGGCAAAACCCAGAC 3139
Qy 3121 AGGAGGCCCAACGACCTCAGCATCGACATCTTACAGCCGAGACACTGTTTCTTGGACGTGCGAG 3180

Qy	1381	CGCCTCAACGGGCACTCTCCGCAAGATCTGTGTGCGAGGACCTGGAGAGCCCGCGAGCC	1440
Db	1429	CGCCTCAACGGCACCTCCGCAAGATCTGTGTGCGAGGACCTGGAGAGCCCGCGAGCC	1488
Qy	1441	ATCGCACTGCAACCGCGTGTATGGGCTCATGTACTGGAACAGACTGGGAGAGAACCTTAAA	1500
Db	1489	ATCGCACTGCAACCGCGTGTATGGGCTCATGTACTGGAACAGACTGGGAGAGAACCTTAAA	1548
Qy	1501	ATCGAGTGTGCAACTTTGGATGGGACGAGCGGCTGTGTGCTCAATGCTCCTCGGG	1560
Db	1549	ATCGAGTGTGCAACTTTGGATGGGACGAGCGGCTGTGTGCTCAATGCTCCTCGGG	1608
Qy	1561	TGGCCCAACGCGCTGGCCCTGGACCTGCAAGAGGGGAGCTTACTGGGAGAGCGCAAG	1620
Db	1609	TGGCCCAACGCGCTGGCCCTGGACCTGCAAGAGGGGAGCTTACTGGGAGAGCGCAAG	1668
Qy	1621	ACAGACAAGATCGAGGTGATCAATGTTGATGGGACGAAGAGCGGACCTCTCGAGGAC	1680
Db	1669	ACAGACAAGATCGAGGTGATCAATGTTGATGGGACGAAGAGCGGACCTCTCGAGGAC	1728
Qy	1681	AAGCTCCGCAATTTTCGGGTTCAAGCTGTGCGGGGACTTCACTACTGGAAGTGTG	1740
Db	1729	AAGCTCCGCAATTTTCGGGTTCAAGCTGTGCGGGGACTTCACTACTGGAAGTGTG	1788
Qy	1741	CAGCGCCGACATCGAGCGGGTGCAAGGTCGAGGTCGAGCGGAGGTCATCATTTGAC	1800
Db	1789	CAGCGCCGACATCGAGCGGGTGCAAGGTCGAGGTCGAGGTCGAGGTCATCATTTGAC	1848
Qy	1801	CAGCTGCCGACCTGATGGGGCTCAAGCTGTGAATGCGCAAGGTCGTGGGACCAAC	1860
Db	1849	CAGCTGCCGACCTGATGGGGCTCAAGCTGTGAATGCGCAAGGTCGTGGGACCAAC	1908
Qy	1861	CCGTGTGCGGACAGAAACGGGGGTGACGCACTGTGCTTCTTCAACCCGACGCAAC	1920
Db	1909	CCGTGTGCGGACAGAAACGGGGGTGACGCACTGTGCTTCTTCAACCCGACGCAAC	1968
Qy	1921	CGGTGTGCTGCCCATCGGCTGGAGCTGTGAGTGACATGAGACCTGTGATGTGCT	1980
Db	1969	CGGTGTGCTGCCCATCGGCTGGAGCTGTGAGTGACATGAGACCTGTGATGTGCT	2028
Qy	1981	GAGGCTTTCTTGGTCTTCAACAGCAGAGCGCGCATCCACAGGATCTCCCTCGAGACCAAT	2040
Db	2029	GAGGCTTTCTTGGTCTTCAACAGCAGAGCGCGCATCCACAGGATCTCCCTCGAGACCAAT	2088
Qy	2041	AACAAACGACGTGGCCATCCCGCTCAACGGCGTCAAGGAGGCTCAGCCCTGGAATTGAT	2100
Db	2089	AACAAACGACGTGGCCATCCCGCTCAACGGCGCTCAAGGAGGCTCAGCCCTGGAATTGAT	2148
Qy	2101	GTGTCCAAACCAACATCTACTGGACAGACGTCAGCCTGAAAGACCATCAGCCGCGCTTC	2160
Db	2149	GTGTCCAAACCAACATCTACTGGACAGACGTCAGCCTGAAAGACCATCAGCCGCGCTTC	2208
Qy	2161	ATGAACGGGAGCTCGGTGGAGCAGCTGTGGAGTTTGGCTTGAATACCCCGAGGGCATG	2220
Db	2209	ATGAACGGGAGCTCGGTGGAGCAGCTGTGGAGTTTGGCTTGAATACCCCGAGGGCATG	2268
Qy	2221	GCCGTTGACTGATGGGCAAGAACTCTACTGGGCGGACAGTGGGACCAACAGAAATCGAA	2280
Db	2269	GCCGTTGACTGATGGGCAAGAACTCTACTGGGCGGACAGTGGGACCAACAGAAATCGAA	2328
Qy	2281	GTGGCGGCGCTGGACGGGCAAGTTCCGGCAAGTCTGTGTGGAGGCACTTGGACAAACCGG	2340
Db	2329	GTGGCGGCGCTGGACGGGCAAGTTCCGGCAAGTCTGTGTGGAGGCACTTGGACAAACCGG	2388
Qy	2341	AGGTGCTGGCCCTGGATCCCAACAAAGGCTCATCTACTGGAACGAGTGGGGCGGCAAG	2400
Db	2389	AGGTGCTGGCCCTGGATCCCAACAAAGGCTCATCTACTGGAACGAGTGGGGCGGCAAG	2448
Qy	2401	CCGAGGATGTCGGGCGCTTCAATGACGGGACCAACTGCATGACGCTGGTGACAGGTTG	2460
Db	2449	CCGAGGATGTCGGGCGCTTCAATGACGGGACCAACTGCATGACGCTGGTGACAGGTTG	2508
Qy	2461	GGCCGGGGCAACGACCTCACCAATTGACTACGTGACCGAGCGGCTCTACTGGACCGACCTG	2520

Db	2509	GGCCGGGCCAAAGACCTCACCATTTGACTACGTGACCGAGCCCTCTACTGGACCGACCTG	2568
Qy	2521	GACACCAACATGATCGAGTGTGCTCAACATGCTGGGTGAGGAGGGGTGCTGATTTGCCGAC	2580
Db	2569	GACACCAACATGATCGAGTGTGCTCAACATGCTGGGTGAGGAGGGGTGCTGATTTGCCGAC	2628
Qy	2581	GATCTCCGCAACCGGTTGCGTCTGAGCCAGTACAGCGATTATATCTACTTGACAGACTGG	2640
Db	2629	GATCTCCGCAACCGGTTGCGTCTGAGCCAGTACAGCGATTATATCTACTTGACAGACTGG	2688
Qy	2641	AATCTGCACAGCATTTGAGCGGGCCGACAAAGACTAGCCGGCCGGAACCCACCTCATCCAG	2700
Db	2689	AATCTGCACAGCATTTGAGCGGGCCGACAAAGACTAGCCGGCCGGAACCCACCTCATCCAG	2748
Qy	2701	GGCCACTGGAATTTGATGAGGACATCTCGGTGTTCATCTCTCCCGCAGGATGGGCTC	2760
Db	2749	GGCCACTGGAATTTGATGAGGACATCTCGGTGTTCATCTCTCCCGCAGGATGGGCTC	2808
Qy	2761	AATGACTGTATGCAACAAAGGGGACGTGTGGGAGCTGTGCTTGGCCATCCCGCGGCGC	2820
Db	2809	AATGACTGTATGCAACAAAGGGGACGTGTGGGAGCTGTGCTTGGCCATCCCGCGGCGC	2868
Qy	2821	CACCGCTGCGGCTGCGCTCACACTACACCTGAGACCCGACGAGCCGCAACTGCAAGCCCG	2880
Db	2869	CACCGCTGCGGCTGCGCTCACACTACACCTGAGACCCGACGAGCCGCAACTGCAAGCCCG	2928
Qy	2881	CCCAACCACTTTGCTGTTGATGAGGACAAATCTGCCATCTAGTCGATGATCCCGGACGAC	2940
Db	2929	CCCAACCACTTTGCTGTTGATGAGGACAAATCTGCCATCTAGTCGATGATCCCGGACGAC	2988
Qy	2941	CAGCAGCGCGGATCTCATCTGCCCCCTGCATGCACTGAGGACGTCAGGACCATCGAC	3000
Db	2989	CAGCAGCGCGGATCTCATCTGCCCCCTGCATGCACTGAGGACGTCAGGACCATCGAC	3048
Qy	3001	TATGACCCACTGGAACAAATTTCTACTTGGGTGATGGGCGCCAGAAACATCAAGCGAGCC	3060
Db	3049	TATGACCCACTGGAACAAATTTCTACTTGGGTGATGGGCGCCAGAAACATCAAGCGAGCC	3108
Qy	3061	AAGGACGAGCGGACCCAGCCCTTTGTTTGTGACCTCTCTGAGCCAGGCAAAACCCAGAC	3120
Db	3109	AAGGACGAGCGGACCCAGCCCTTTGTTTGTGACCTCTCTGAGCCAGGCAAAACCCAGAC	3168
Qy	3121	AGGACGAGCCCAACGACCTCAGCATCGACATCTACAGCCGACACATGTTCTTGACGTCGAG	3180
Db	3169	AGGACGAGCCCAACGACCTCAGCATCGACATCTACAGCCGACACATGTTCTTGACGTCGAG	3228
Qy	3181	GCCACCAATATCATCAACGCTCCACAGGCTGAGCGGGGAAAGCCATGGGGGTGTTGCTGGT	3240
Db	3229	GCCACCAATATCATCAACGCTCCACAGGCTGAGCGGGGAAAGCCATGGGGGTGTTGCTGGT	3288
Qy	3241	GGGACCGCGACAGCCCGAGGCGCATCGTGTCAACGGGAGCGAGGTTACTGTACTTC	3300
Db	3289	GGGACCGCGACAGCCCGAGGCGCATCGTGTCAACGGGAGCGAGGTTACTGTACTTC	3348
Qy	3301	ACCAACATGACGAGCCGCGGACGCAAGATCGAAACGCGCAGCCCTGGAAGCCAGAGCGC	3360
Db	3349	ACCAACATGACGAGCCGCGGACGCAAGATCGAAACGCGCAGCCCTGGAAGCCAGAGCGC	3408
Qy	3361	GAGGTCTCTTTCACCAACCGGCTCATCGCCCTGTGGCCCTGTGGTGGGACAAACACTG	3420
Db	3409	GAGGTCTCTTTCACCAACCGGCTCATCGCCCTGTGGCCCTGTGGTGGGACAAACACTG	3468
Qy	3421	GGCAAGCTGTTCTGGGTGGACCGGACCTGAAGCGCATTTGAGACTGTGACTGTGACGG	3480
Db	3469	GGCAAGCTGTTCTGGGTGGACCGGACCTGAAGCGCATTTGAGACTGTGACTGTGACGG	3528
Qy	3481	GCCAAACCGCTGACCTCGGAGGACGCAACATCTGTGAGCCCTCTGGGCGCTGACCACTCTT	3540
Db	3529	GCCAAACCGCTGACCTCGGAGGACGCAACATCTGTGAGCCCTCTGGGCGCTGACCACTCTT	3588
Qy	3541	GGCAAGCATCTCTACTGGATCGAACCGCGCAGCAGCATGATCGAGCGTGTGGAGAGAC	3600

Db 3589 GCGAAGCATCTTACTGGATCGACCGCCAGCAGCAGATGATCGAGCGTGTGGAGAAGACC 3648
Qy 3601 ACCGGGGAACAGCGACTCGCATCCAGGCGCGTGTGCGCCACCTCAGCTGGCATCCATGCA 3660
Db 3649 ACCGGGGAACAGCGACTCGCATCCAGGCGCGTGTGCGCCACCTCAGCTGGCATCCATGCA 3708
Qy 3661 GTGGAGGAAGTCAGCTCGAGGAGTTCTCAGCCCAACCATGTGCCCGTGACAATGGTGGC 3720
Db 3709 GTGGAGGAAGTCAGCTCGAGGAGTTCTCAGCCCAACCATGTGCCCGTGACAATGGTGGC 3768
Qy 3721 TGCTCCCAATCTGTATTGCCAAGGGTGATGGGACACACGGTGTCTATGCCAGTCAC 3780
Db 3769 TGCTCCCAATCTGTATTGCCAAGGGTGATGGGACACACGGTGTCTATGCCAGTCAC 3828
Qy 3781 CTGCTGCTCCGCGAAGCTGTGACCTGTGAGAGCGCCACCTGTCTCCCGGACACAG 3840
Db 3829 CTGCTGCTCCGCGAAGCTGTGACCTGTGAGAGCGCCACCTGTCTCCCGGACACAG 3888
Qy 3841 TTGTGATGTGCACAGGGAGATCGACTGTATCCCGGGGCTGGCGCTGTGACGGCTTT 3900
Db 3889 TTGTGATGTGCACAGGGAGATCGACTGTATCCCGGGGCTGGCGCTGTGACGGCTTT 3948
Qy 3901 CCGAGTGGATGACACAGAGCAGCAGAGGCTGCCCGCTGTGCTCCCGCGCCAGTTC 3960
Db 3949 CCGAGTGGATGACACAGAGCAGCAGAGGCTGCCCGCTGTGCTCCCGCGCCAGTTC 4008
Qy 3961 CCCTGCGCGCGGGTCACTGTGTGACCTGTGCGCTGTGCGAGCGGAGGAGACTGT 4020
Db 4009 CCCTGCGCGCGGGTCACTGTGTGACCTGTGCGCTGTGCGAGCGGAGGAGACTGT 4068
Qy 4021 CAGGACCGCTCAGAGCGGAGACTGTGACGCGCATCTGCTGCCCAACAGTTCGGGTGT 4080
Db 4069 CAGGACCGCTCAGAGCGGAGACTGTGACGCGCATCTGCTGCCCAACAGTTCGGGTGT 4128
Qy 4081 CGGAGCGCGAGTGTCTCATCAACAGCAGTGGCATCTCTTCCCGCGCATGTATCGAC 4140
Db 4129 CGGAGCGCGAGTGTCTCATCAACAGCAGTGGCATCTCTTCCCGCGCATGTATCGAC 4188
Qy 4141 GGCTCCGAGCAGCTCATGTGTGAATACCAAGCGCGCTCAGAGCAGCGCGCCAC 4200
Db 4189 GGCTCCGAGCAGCTCATGTGTGAATACCAAGCGCGCTCAGAGCAGCGCGCCAC 4248
Qy 4201 AGCAGTGCATCGGCGCGCTCAITGGCATCATCTCTCTCTCTGTCATGGGTGGTGC 4260
Db 4249 AGCAGTGCATCGGCGCGCTCAITGGCATCATCTCTCTCTCTGTCATGGGTGGTGC 4308
Qy 4261 TATTTTGTGCCAGCGGTGGTGTGCGAGCTATGCGGGGCGCAACGGGCGCTTCCG 4320
Db 4309 TATTTTGTGCCAGCGGTGGTGTGCGAGCTATGCGGGGCGCAACGGGCGCTTCCG 4368
Qy 4321 CACGAGTATGTCAGCGGACCCCGCAGCTGCGCTCAATTTTCATAGCCCGGGCGTTCC 4380
Db 4369 CACGAGTATGTCAGCGGACCCCGCAGCTGCGCTCAATTTTCATAGCCCGGGCGTTCC 4428
Qy 4381 CAGCATGGCCCCCTTCACAGGCATCGCATGCGGAAAGTCCATGATGAGCTCCGTCAGCGCTG 4440
Db 4429 CAGCATGGCCCCCTTCACAGGCATCGCATGCGGAAAGTCCATGATGAGCTCCGTCAGCGCTG 4488
Qy 4441 ATGGGGGCGCGGGCGGGTGCCTCTGTACACGGAACACGTCACAGGGGCTCGTTC 4500
Db 4489 ATGGGGGCGCGGGCGGGTGCCTCTGTACACGGAACACGTCACAGGGGCTCGTTC 4548
Qy 4501 AGCAGCTGTCACGAGCGGACCGCTGTACCGCGGATCCCTGACCCCGCGCCCTCC 4560
Db 4549 AGCAGCTGTCACGAGCGGACCGCTGTACCGCGGATCCCTGACCCCGCGCCCTCC 4608
Qy 4561 CCGGCCACGAGCCCTCCCTGTACAAACATGACATGTTTCTACTCTTCAAAATTCGCGCC 4620
Db 4609 CCGGCCACGAGCCCTCCCTGTACAAACATGACATGTTTCTACTCTTCAAAATTCGCGCC 4668
Qy 4621 ACTGCGAGACCGTACAGGCCCTACATCATTCAGGAATGGCGCCCGCGACGACCGCTTCG 4680
Db 4669 ACTGCGAGACCGTACAGGCCCTACATCATTCAGGAATGGCGCCCGCGACGACCGCTTCG 4728

Qy 4681 AGCACCGACGTGTGTGACAGCGACTACAGCGCCAGCGCTGGAAGGCCAGCAAGTACTAC 4740
Db 4729 AGCACCGACGTGTGTGACAGCGACTACAGCGCCAGCGCTGGAAGGCCAGCAAGTACTAC 4788
Qy 4741 CTGGATTTGAACCTCGGACTCAGACCCCTTATCCACCCACCCACCCACCCAGCCAGTAC 4800
Db 4789 CTGGATTTGAACCTCGGACTCAGACCCCTTATCCACCCACCCACCCAGCCAGTAC 4848
Qy 4801 CTGTCCGGGAGGACAGCTGCCCGCTTCCCGCCGACCGGACCGAGAGGAGTACTTCCATCTC 4860
Db 4849 CTGTCCGGGAGGACAGCTGCCCGCTTCCCGCCGACCGGACCGAGAGGAGTACTTCCATCTC 4908
Qy 4861 TTCCCGCCCTCCGCTGCCCTGACCGACTCATCTGACCTCGCGCGGCGGACCTCTGGC 4920
Db 4909 TTCCCGCCCTCCGCTGCCCTGACCGACTCATCTGACCTCGCGCGGCGGACCTCTGGC 4968
Qy 4921 TTCTCTGTGCCCCCTGTAAATAGTTTTAAATATGAACAAGAAAAAATATATATTTATGAT 4980
Db 4969 TTCTCTGTGCCCCCTGTAAATAGTTTTAAATATGAACAAGAAAAAATATATTTATGAT 5028
Qy 4981 TTAAAAATAAATAATATTGGATTTTAAAAACATGAGAAATGTGAATGTGATGGGTG 5040
Db 5029 TTAAAAATAAATAATATTGGATTTTAAAAACATGAGAAATGTGAATGTGATGGGTG 5088
Qy 5041 GGCAGGGCTGGGAGAACTTTGTA 5063
Db 5089 GGCAGGGCTGGGAGAACTTTGTA 5111

RESULT 11

ADB98797
ID ADB98797 standard; DNA; 5278 BP.

XX ADB98797;

XX 04-DEC-2003 (first entry)

XX Human ZMAXGI-3AS DNA sequence.

XX Osteopathic; Gene therapy; High Bone Mass; HBM; LRP5; Zmax1; LRP6;

XX bone mass modulation; osteoporosis; human; ds.

XX Homo sapiens.

XX WO200292000-A2.

XX 21-NOV-2002.

XX 13-MAY-2002; 2002WO-US014877.

XX 11-MAY-2001; 2001US-0290071P.

XX 17-MAY-2001; 2001US-0291311P.

XX 01-FEB-2002; 2002US-0353058P.

XX 04-MAR-2002; 2002US-0361293P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX (AMHP) WYETH.

XX Allen K, Anisowicz A, Graham JR, Morales A, Yaworsky PJ, Liu W;

XX WPI; 2003-129214/12.

XX New nucleic acid comprising a mutation in LRP5 or LRP6, useful for

XX diagnosing a HBM-like phenotype in a subject and for preparing a

XX composition for modulating bone mass and/or lipid levels in a subject

XX suffering from e.g. osteoporosis.

XX Disclosure; Fig 25; 629pp; English.

XX The present invention relates to High Bone Mass (HBM), LRP5 (Zmax1) and

XX LRP6 mutants, which results in a HBM-like phenotype when expressed in a

XX cell. The HBM-like phenotype results in bone mass modulation and/or lipid

QY 2284 GCGGGCTGGACGGGAGTTCCGGCAGTCTCGTGTGGAGGAGCTTGGACAAACCCGAGG 2343
DB 2281 GCGGGCTGGACGGGAGTTCCGGCAGTCTCGTGTGGAGGAGCTTGGACAAACCCGAGG 2340
QY 2344 TCGTGTGGCCCTGGATCCCAACAAAGGGCTACACTTACTTGGACCGAGTGGGGCGGCAAGCCG 2403
DB 2341 TCGTGTGGCCCTGGATCCCAACAAAGGGCTACACTTACTTGGACCGAGTGGGGCGGCAAGCCG 2400
QY 2404 AGGATCTGTGGCCCTTTCATGGACGGGACCAACTGTGATGACGCTGGTGGACAAAGTGGGC 2463
DB 2401 AGGATCTGTGGCCCTTTCATGGACGGGACCAACTGTGATGACGCTGGTGGACAAAGTGGGC 2460
QY 2464 CGGGCCACGACCTCACCATTTAGCTTACGCTGACGCGCCCTCTACTTGGACCGACCTGGAC 2523
DB 2461 CGGGCCACGACCTCACCATTTAGCTTACGCTGACGCGCCCTCTACTTGGACCGACCTGGAC 2520
QY 2524 ACCAACATGATCGAGTCTCCAAACATGTGGGTGAGGAGCGGTCTGATTTGCCGACGAT 2583
DB 2521 ACCAACATGATCGAGTCTCCAAACATGTGGGTGAGGAGCGGTCTGATTTGCCGACGAT 2580
QY 2584 CTCGCGCACCCGTTTCGCTCTGACGAGTACAGCGATTTATATCTACTTGGACAGACTGGAAAT 2643
DB 2581 CTCGCGCACCCGTTTCGCTCTGACGAGTACAGCGATTTATATCTACTTGGACAGACTGGAAAT 2640
QY 2644 CTGCACAGCATTTGAGCGGGCGGACAAAGACTAGCGCGGGAACCGCACCTCATCCAGGGC 2703
DB 2641 CTGCACAGCATTTGAGCGGGCGGACAAAGACTAGCGCGGGAACCGCACCTCATCCAGGGC 2700
QY 2704 CACCTGGACTTCTGATGGACATCTCTGTGTGTTCACCTCTCCCGCAGGATGGCTCAAT 2763
DB 2701 CACCTGGACTTCTGATGGACATCTCTGTGTGTTCACCTCTCCCGCAGGATGGCTCAAT 2760
QY 2764 GACTGTATGCAACAACGGGGAGTGTGGGAGCTGTGCTTGGCATCCCGGGCGGCAC 2823
DB 2761 GACTGTATGCAACAACGGGGAGTGTGGGAGCTGTGCTTGGCATCCCGGGCGGCAC 2820
QY 2824 CGCTGGGCTGGCTCACACTACACCTTGACCCCGACGCGCAACTGACAGCCGCCC 2883
DB 2821 CGCTGGGCTGGCTCACACTACACCTTGACCCCGACGCGCAACTGACAGCCGCCC 2880
QY 2884 ACCACCTTCTTGTCTCAGCCAGAAATCTGCCATCAGTGGATGATCCCGGACGACAG 2943
DB 2881 ACCACCTTCTTGTCTCAGCCAGAAATCTGCCATCAGTGGATGATCCCGGACGACAG 2940
QY 2944 CACAGCCGGATCTCATCTGCTCCCTGCATGGAATGAGGAGAGTCAAGGCCATCGACTAT 3003
DB 2941 CACAGCCGGATCTCATCTGCTCCCTGCATGGAATGAGGAGAGTCAAGGCCATCGACTAT 3000
QY 3004 GACCCACTGGACAAAGTTTCATCTACTGGGTGGATGGGCGCCAGAACATCAAGCGAGCCAAAG 3063
DB 3001 GACCCACTGGACAAAGTTTCATCTACTGGGTGGATGGGCGCCAGAACATCAAGCGAGCCAAAG 3060
QY 3064 GACGACGGGACCCAGCCCTTTGTTTGTGACCTCTCTGAGCCCAAGGCCAAAACCCAGACAGG 3123
DB 3061 GACGACGGGACCCAGCCCTTTGTTTGTGACCTCTCTGAGCCCAAGGCCAAAACCCAGACAGG 3120
QY 3124 CAGCCCCACGACCTCAGCATCGACATCTACAGCCGGAACATGTTTCTGACGTCGGAGGCC 3183
DB 3121 CAGCCCCACGACCTCAGCATCGACATCTACAGCCGGAACATGTTTCTGACGTCGGAGGCC 3180
QY 3184 ACCAATACATCAACGTCACAGGCTGACGGGGGAGCCATGGGGGTGGTCTGGTGGG 3243
DB 3181 ACCAATACATCAACGTCACAGGCTGACGGGGGAGCCATGGGGGTGGTCTGGTGGG 3240
QY 3244 GACCCGCAACAGCCCAAGGCCCATCGTCTCAACGCGGAGCGAGGGTACTGTTACTTTCACC 3303
DB 3241 GACCCGCAACAGCCCAAGGCCCATCGTCTCAACGCGGAGCGAGGGTACTGTTACTTTCACC 3300
QY 3304 AACATGACGAGCCGGGACGCAAGATCGAAGCGCAGCCCTTGGAGCGGACCGAGCGCGAG 3363
DB 3301 AACATGACGAGCCGGGACGCAAGATCGAAGCGCAGCCCTTGGAGCGGACCGAGCGCGAG 3360
QY 3364 GTCTCTTTACACCCGGGCTCATCCGCCCTGTGGCCCTGTGGTGGACACACACTGGGC 3423

DB 3361 GTCTCTTTACACCCGGGCTCATCCGCCCTGTGGCCCTGTGGTGGACACACACTGGGC 3420
QY 3424 AAGCTGTTCTGGGTGGACGGGACCTGAAAGCGCATTTGAGAGCTGTGACCTGTGTCAAGGGCC 3483
DB 3421 AAGCTGTTCTGGGTGGACGGGACCTGAAAGCGCATTTGAGAGCTGTGACCTGTGTCAAGGGCC 3480
QY 3484 AACCGCTGACCTTGGAGGAGCGCAACATCGTGCAGCTCTGGGCTGTGACCATCTTGGC 3543
DB 3481 AACCGCTGACCTTGGAGGAGCGCAACATCGTGCAGCTCTGGGCTGTGACCATCTTGGC 3540
QY 3544 AAGCATCTCTACTTGGATCGACCGCCAGCAGCAGATGATCGAGCGTGTGGAGAACACCAAC 3603
DB 3541 AAGCATCTCTACTTGGATCGACCGCCAGCAGCAGATGATCGAGCGTGTGGAGAACACCAAC 3600
QY 3604 GGGGACAAAGCGGACTCTGCATCCAGGGCGGTGTGCGCCACCTCTCATCTGGCATCATGCAAGT 3663
DB 3601 GGGGACAAAGCGGACTCTGCATCCAGGGCGGTGTGCGCCACCTCTCATCTGGCATCATGCAAGT 3660
QY 3664 GAGGAAGTCAAGCTTGGAGGAGTCTCAGCCACCCATGTGCGCCGTGCAATGTGGTGGCTGC 3723
DB 3661 GAGGAAGTCAAGCTTGGAGGAGTCTCAGCCACCCATGTGCGCCGTGCAATGTGGTGGCTGC 3720
QY 3724 TCCCAACATCTGTATTGTCAAGGGGTGATGGGACACACCGGTCTCATGCCCCAGTCCACCTC 3783
DB 3721 TCCCAACATCTGTATTGCCAAGGGGTGATGGGACACACCGGTCTCATGCCCCAGTCCACCTC 3780
QY 3784 GTGCTCTGACAAACCTGTGTGACCTGTGGAGAGCGGCCACCTGCTCCCGGACGAGTTT 3843
DB 3781 GTGCTCTGACAAACCTGTGTGACCTGTGGAGAGCGGCCACCTGCTCCCGGACGAGTTT 3840
QY 3844 GCATGTGCCACAGGGGAGATCGACTGTATCCCGGGGCTGTGAGCGCTTTCC 3903
DB 3841 GCATGTGCCACAGGGGAGATCGACTGTATCCCGGGGCTGTGAGCGCTTTCC 3900
QY 3904 GAGTGCATGACACAGAGCGAGAGGGTGCCTGTGTCTCCCGGCGCAGTTCC 3963
DB 3901 GAGTGCATGACACAGAGCGAGAGGGTGCCTGTGTCTCCCGGCGCAGTTCC 3960
QY 3964 TGCGCGGGGTGAGTGTGGACCTGCGCTGTGCTGCGAGCGGAGGAGTGTGAG 4023
DB 3961 TGCGCGGGGTGAGTGTGGACCTGCGCTGTGCTGCGAGCGGAGGAGTGTGAG 4020
QY 4024 GACCGCTCAGACGAGCGGAGTGTGAGCGCATCTGCTGCGCCAAACAGTTCCGGTGTGCG 4083
DB 4021 GACCGCTCAGACGAGCGGAGTGTGAGCGCATCTGCTGCGCCAAACAGTTCCGGTGTGCG 4080
QY 4084 AGCGGCGAGTGTGTCTCATCAAAACAGAGTGCAGCTCTTCCCGGACTGTATGACGGC 4143
DB 4081 AGCGGCGAGTGTGTCTCATCAAAACAGAGTGCAGCTCTTCCCGGACTGTATGACGGC 4140
QY 4144 TCCGACGAGTCTGATGTGAATCAACAGCGCCCTCAGACGACAGCCCGGCCACAGC 4203
DB 4141 TCCGACGAGTCTGATGTGAATCAACAGCGCCCTCAGACGACAGCCCGGCCACAGC 4200
QY 4204 AGTGCATCGGGCCCGCTCATTTGGCATCATCTCTCTCTCTTCTGTCATGGGTGTCTAT 4263
DB 4201 AGTGCATCGGGCCCGCTCATTTGGCATCATCTCTCTCTCTCTCTGTCATGGGTGTCTAT 4260
QY 4264 TTTGTGTGCCAGCGGTGTGTCAGCGCTATGCGGGGCGCAACGGGCGCTTCCCGGAC 4323
DB 4261 TTTGTGTGCCAGCGGTGTGTCAGCGCTATGCGGGGCGCAACGGGCGCTTCCCGGAC 4320
QY 4324 GAGTATGTACGGGAGCCCGCAGTGCCTCTCAATTTCTAGCCCCGGGGGTTCCCGAG 4383
DB 4321 GAGTATGTACGGGAGCCCGCAGTGCCTCTCAATTTCTAGCCCCGGGGGTTCCCGAG 4380
QY 4384 CATGGCCCTTTACAGGAGCATCGCATGCGGAAAGTCCATGATGAGCTCCGTGAGCGTGTG 4443
DB 4381 CATGGCCCTTTACAGGAGCATCGCATGCGGAAAGTCCATGATGAGCTCCGTGAGCGTGTG 4440
QY 4444 GGGGGCGGGGGGGTGGCTGTGTAAGACCGGAAACGATCAAGGGGCTCTGTCCAGC 4503

Db	560		TACATGTA	CTGGACAGAC	CTGGGTTGAGACGCCCCGGAAATTGAGCGGGCAGGGATGGAATGGC	619
Qy	601	AGCACCCGGGAAGATCA	TTGTGGACTCTGGGACATTTTACTTGGCCCAATGCACTGACCATCGAC	660		
Db	620	AGCACCCGGGAAGATCA	TTGTGGACTCTGGGACATTTTACTTGGCCCAATGCACTGACCATCGAC	679		
Qy	661	CTGAGGAGGACGAAGACTCTA	CTGGGCTGACCGCAAGCTCAGCTTCTATCCACCGTGCACAC	720		
Db	680	CTGAGGAGGACGAAGACTCTA	CTGGGCTGACCGCAAGCTCAGCTTCTATCCACCGTGCACAC	739		
Qy	721	CTGACAGGCTCGTTCCGGCAGAGTG	TGTGGAGGGCAGCCTGACGACACCCCTTCGCCCTG	780		
Db	740	CTGARCGGCTCGTTCCGGCAGAGTG	TGTGGAGGGCAGCCTGACGACACCCCTTCGCCCTG	799		
Qy	781	ACGCTCTCCGGGACACTCTGT	TACTGGACAGACTGGCAGACCCGCTCCATCCATGCTCTGC	840		
Db	800	ACGCTCTCCGGGACACTCTGT	TACTGGACAGACTGGCAGACCCGCTCCATCCATGCTCTGC	859		
Qy	841	AACAAGCGCACTGGGGGGAAGAGAA	GGGAGTCTGTAGTGCCCTTACTACCCATGAGGAGGACAAT	900		
Db	860	AACAAGCGCACTGGGGGGAAGAGAG	ATCTGTAGTGCCCTTACTACCCATGAGGAGGACAAT	919		
Qy	901	ATCCAGGTGTGCTGAGCCAGGACGGC	AGCCTTTCTTCCACACTCGCTCTGTGAGGAGGACAAT	960		
Db	920	ATCCAGGTGTGCTGAGCCAGGACGGC	AGCCTTTCTTCCACACTCGCTCTGTGAGGAGGACAAT	979		
Qy	961	GGGGCTGTCCCACTGTGCTCTGCT	GTCCCAAGCGAGCCTTTCTTACACATGCGCCTGC	1020		
Db	980	GGGGCTGTCTCCCACTGTGCTCTG	CTGCCCAAGCGAGCCTTTCTTACACATGCGCCTGC	1039		
Qy	1021	CCCAACGGTGTGACGTGTCAGGACAA	CGGCAGGACGTGTAAAGCAGAGCCGAGGAGGTG	1080		
Db	1040	CCCAACGGTGTGACGTGTCAGGACAA	CGGCAGGACGTGTAAAGCAGAGCCGAGGAGGTG	1099		
Qy	1081	CTGTGCTGCGCCCGCGCAGCGAAC	CTTACGGAGATCTCGCTGGACACGCCCGGACTTCAACC	1140		
Db	1100	CTGTGCTGCGCCCGCGCAGCGAAC	CTTACGGAGATCTCGCTGGACACGCCCGGACTTCAACC	1159		
Qy	1141	GACATCGTGTGCAAGTGGACGACAT	CCGGCAGCCCATTCGCCATCGACTACGACCCGCTA	1200		
Db	1160	GACATCGTGTGCAAGTGGACGACAT	CCGGCAGCCCATTCGCCATCGACTACGACCCGCTA	1219		
Qy	1201	GAGGGCTATGTCTACTGGACAGATG	ACGAGGTGGCGGCCCATTCGGCAGGGCGGTACTCTGGAC	1260		
Db	1220	GAGGGCTATGTCTACTGGACAGATG	ACGAGGTGGCGGCCCATTCGGCAGGGCGGTACTCTGGAC	1279		
Qy	1261	GGTCTTGGGGCGACAGCTGTGTTCAA	CCCGAGATCAACGACCCCGATGGCATGCGGTC	1320		
Db	1280	GGTCTTGGGGCGACAGCTGTGTTCAA	CCCGAGATCAACGACCCCGATGGCATGCGGTC	1339		
Qy	1321	GACTGGGTGGCCCGAAACCTCTACT	TGGACCGACACACGGGCGACGACCCGATCGAGGTGACG	1380		
Db	1340	GACTGGGTGGCCCGAAACCTCTACT	TGGACCGACACACGGGCGACGACCCGATCGAGGTGACG	1399		
Qy	1381	CGCTCTAACGGGCACTCCCGCAAGAT	CTTGGTGTGCGAGGACCTTGGACGAGCCCCCGAGCC	1440		
Db	1400	CGCTCTAACGGGCACTCCCGCAAGAT	CTTGGTGTGCGAGGACCTTGGACGAGCCCCCGAGCC	1459		
Qy	1441	ATCGCACTGCACCCCGTGATGGGCT	CTATGTACTTGGACAGACTGGGGAGAGAACCTTAAA	1500		
Db	1460	ATCGCACTGCACCCCGTGATGGGCT	CTATGTACTTGGACAGACTGGGGAGAGAACCTTAAA	1519		
Qy	1501	ATCGAGTGTGCCAACTTTGGATGGC	AGGACGGCGTGTGTGTTCAATGCTCTCCCTCGGG	1560		
Db	1520	ATCGAGTGTGCCAACTTTGGATGGC	AGGACGGCGTGTGTGTTCAATGCTCTCCCTCGGG	1579		
Qy	1561	TGGCCCAACGGCTGGCCCTGGACCTT	GCAGGAGGGGAGCTTACTTGGGAGAGCGCCAG	1620		
Db	1580	TGGCCCAACGGCTGGCCCTGGACCTT	GCAGGAGGGGAGCTTACTTGGGAGAGCGCCAG	1639		
Qy	1621	ACAGACAAGATTCAGGTGTGATCAAT	GTGTTGATGGGACGAAGAGGGGGAACCTTCTCTGGAGAC	1680		

QY 2761 AATGACTGTATGCAACAACGGGAGTGTGGCAGCTGTGCTTGCCATCCCGGCGGC 2820
DB |||||
QY 2780 AATGACTGTATGCAACAACAAGGGCAGTGTGGCAGCTGTGCTTGCCATCCCGGCGGC 2839
DB |||||
QY 2821 CACCGCTGGGTGGCTCTCACTACACCTTGGACCCAGAGCGGCAACTGACAGCCG 2880
DB |||||
QY 2840 CACCGCTGGGTGGCTCTCACTACACCTTGGACCCAGAGCGGCAACTGACAGCCG 2899
DB |||||
QY 2881 CCCACACTTCTTCTGCTGTAGCCAGAAATCTGCCATCAGTCGATGATCCCGACGAC 2940
DB |||||
QY 2900 CCCACACTTCTTCTGCTGTAGCCAGAAATCTGCCATCAGTCGATGATCCCGACGAC 2959
DB |||||
QY 2941 CAGCACGCCCGGATCTCATCTGCTGCTTGGATGAGTGGAGAACTGAGAACTGAGAACTGAG 3000
DB |||||
QY 2960 CAGCACGCCCGGATCTCATCTGCTGCTTGGATGAGTGGAGAACTGAGAACTGAGAACTGAG 3019
DB |||||
QY 3001 TATGACCACTGGACAAAGTTGATCTACTGGGTGGATGGGCGCAGAACTACAGCGAGCC 3060
DB |||||
QY 3020 TATGACCACTGGACAAAGTTGATCTACTGGGTGGATGGGCGCAGAACTACAGCGAGCC 3079
DB |||||
QY 3061 AAGGACGAGCGGACCCAGCCCTTGTGTTGACCTCTGAGCCAGGCGCAAAACCCAGAC 3120
DB |||||
QY 3080 AAGGACGAGCGGACCCAGCCCTTGTGTTGACCTCTGAGCCAGGCGCAAAACCCAGAC 3139
DB |||||
QY 3121 AAGGACGAGCGGACCCAGCCCTTGTGTTGACCTCTGAGCCAGGCGCAAAACCCAGAC 3180
DB |||||
QY 3140 AAGGACGAGCGGACCCAGCCCTTGTGTTGACCTCTGAGCCAGGCGCAAAACCCAGAC 3199
DB |||||
QY 3181 GCCACCAATACATCAACGTGCAAGGCTGAGCGGGGAAGCCATGGGGGTGGTGTGCT 3240
DB |||||
QY 3200 GCCACCAATACATCAACGTGCAAGGCTGAGCGGGGAAGCCATGGGGGTGGTGTGCT 3259
DB |||||
QY 3241 GGGGACCGCGACAGCCAGGCGCCATCTGCTGCTGAGCGGAGCGGCTGCTGCTGCT 3300
DB |||||
QY 3260 GGGGACCGCGACAGCCAGGCGCCATCTGCTGCTGAGCGGAGCGGCTGCTGCTGCT 3319
DB |||||
QY 3301 ACCAACTGAGGACCGGGGACGCAAGATCGAAACGCGCAGCCCTGGAACGCGCAGCGC 3360
DB |||||
QY 3320 ACCAACTGAGGACCGGGGACGCAAGATCGAAACGCGCAGCCCTGGAACGCGCAGCGC 3379
DB |||||
QY 3361 GAGGTCTCTTCAACACCGGCTCATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3420
DB |||||
QY 3380 GAGGTCTCTTCAACACCGGCTCATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3439
DB |||||
QY 3421 GCGAAGCTGTTCTGGGTGGACGCGGACCTGAGCGCATTGAGAGCTGAGCTGCTGAGG 3480
DB |||||
QY 3440 GCGAAGCTGTTCTGGGTGGACGCGGACCTGAGCGCATTGAGAGCTGAGCTGCTGAGG 3499
DB |||||
QY 3481 GCCAACCGCTGACCCCTGGAGGACGCGCAACATCGTGACGCTCTGGGCTGACCATCTT 3540
DB |||||
QY 3500 GCCAACCGCTGACCCCTGGAGGACGCGCAACATCGTGACGCTCTGGGCTGACCATCTT 3559
DB |||||
QY 3541 GCGAAGCATCTCTACTGATCGAACCGCAGCAGCAGATGATGAGCGGTGGAGAGACC 3600
DB |||||
QY 3560 GCGAAGCATCTCTACTGATCGAACCGCAGCAGCAGATGATGAGCGGTGGAGAGACC 3619
DB |||||
QY 3601 ACCGGGACAGCGGACCTCGCATCCAGGCGCTGCTGCGCCACCTCACTGGCATCATGCA 3660
DB |||||
QY 3620 ACCGGGACAGCGGACCTCGCATCCAGGCGCTGCTGCGCCACCTCACTGGCATCATGCA 3679
DB |||||
QY 3661 GTGAGGAAAGTCAGCCTGGAGGAGTTCTCAGCCCAACCATGTGCCGTGACAAATGGTGGC 3720
DB |||||
QY 3680 GTGAGGAAAGTCAGCCTGGAGGAGTTCTCAGCCCAACCATGTGCCGTGACAAATGGTGGC 3739
DB |||||
QY 3721 TGCTCCCACTCTGATGTCAGAGGTGATGGGACACCAACCGGTGCTCATGCCAGTCCAC 3780
DB |||||
QY 3740 TGCTCCCACTCTGATGTCAGAGGTGATGGGACACCAACCGGTGCTCATGCCAGTCCAC 3799
DB |||||
QY 3781 CTGCTGCTCTGCAAGACCTGCTGACCTGTGGAGAGCGCCACCTGCTCCCGGACGAC 3840
DB |||||
QY 3800 CTGCTGCTCTGCAAGACCTGCTGACCTGTGGAGAGCGCCACCTGCTCCCGGACGAC 3859
DB |||||

QY 3841 TTTGCTATGTGCCACAGAGGAGATCGACTGTATCCCGGGGCTGTGCGCTGTGACGGCTTT 3900
DB |||||
QY 3860 TTTGCTATGTGCCACAGAGGAGATCGACTGTATCCCGGGGCTGTGCGCTGTGACGGCTTT 3919
DB |||||
QY 3901 CCCAGTGCATGATGACAGAGCGACGAGAGGGCTGCCCCGTGTGCTCCGCGCCGCTTCC 3960
DB |||||
QY 3920 CCCAGTGCATGATGACAGAGCGACGAGAGGGCTGCCCCGTGTGCTCCGCGCCGCTTCC 3979
DB |||||
QY 3961 CCCTGCGCGGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4020
DB |||||
QY 3980 CCCTGCGCGGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4039
DB |||||
QY 4021 CAGGACCGCTCAGACGAGCGGACTGTGACGCCATCTGCTGCTGCTGCTGCTGCTGCTGCT 4080
DB |||||
QY 4040 CAGGACCGCTCAGACGAGGCTGCTGACGCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4099
DB |||||
QY 4081 GCGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4140
DB |||||
QY 4100 GCGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4159
DB |||||
QY 4141 GCTCCGACGAGCTCATGTGTGAAATCAACAAGCGCCCTCAGACGACAGCGCGGCCAC 4200
DB |||||
QY 4160 GCTCCGACGAGCTCATGTGTGAAATCAACAAGCGCCCTCAGACGACAGCGCGGCCAC 4219
DB |||||
QY 4201 AGCAGTGCATCGGGCGGCTCATTTGGCATCATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4260
DB |||||
QY 4220 AGCAGTGCATCGGGCGGCTCATTTGGCATCATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4279
DB |||||
QY 4261 TATTTTGTGCTGACAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4320
DB |||||
QY 4280 TATTTTGTGCTGACAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4339
DB |||||
QY 4321 CACGAGTATGTACGCGGACCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4380
DB |||||
QY 4340 CACGAGTATGTACGCGGACCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4399
DB |||||
QY 4381 CAGCATGGCCCCCTTCAAGGATCGCATGCGGAAAGTCCATGATGAGCTCGCTGAGCCTG 4440
DB |||||
QY 4400 CAGCATGGCCCCCTTCAAGGATCGCATGCGGAAAGTCCATGATGAGCTCGCTGAGCCTG 4459
DB |||||
QY 4441 ATGGGGGCGGGGCGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4500
DB |||||
QY 4460 ATGGGGGCGGGGCGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4519
DB |||||
QY 4501 AGCAGCTGCTCAGACGAGGCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4560
DB |||||
QY 4520 AGCAGCTGCTCAGACGAGGCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4579
DB |||||
QY 4561 CCGGCGCAGGACCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4620
DB |||||
QY 4580 CCGGCGCAGGACCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4639
DB |||||
QY 4621 ACTCGGAGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4680
DB |||||
QY 4640 ACTCGGAGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4699
DB |||||
QY 4681 AGCACCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4740
DB |||||
QY 4700 AGCACCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4759
DB |||||
QY 4741 CTGATTTGAACTCGGACTCAGACCCCTATCCACCCCGCCAGCGCCCGCCAGCGGCTGAC 4800
DB |||||
QY 4760 CTGATTTGAACTCGGACTCAGACCCCTATCCACCCCGCCAGCGCCCGCCAGCGGCTGAC 4819
DB |||||
QY 4801 CTGCTCGGAGGACGAGTGGCGCCCTCTGCGCCCGCCAGCGGAGAGCTACTTCCATCTC 4860
DB |||||
QY 4820 CTGCTCGGAGGACGAGTGGCGCCCTCTGCGCCCGCCAGCGGAGAGCTACTTCCATCTC 4879
DB |||||
QY 4861 TTCCCGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4920
DB |||||
QY 4880 TTCCCGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4939
DB |||||
QY 4921 TTCTCTGTGCCCCCTGTAATAGTTTTTAAATATGAACAAAGAAAAAATATATTTTATGAT 4980
DB |||||

QY 1561 TGGCCCAACGGGCTGGCCCTGGACCTGACGAGGGGAGCTTACTTGGGGAGACGCCAAG 1620
Db |||||
1580 TGGCCCAACGGGCTGGCCCTGGACCTGACGAGGGGAGAGCTTACTTGGGGAGACGCCAAG 1639
QY |||||
1621 ACAGACAAGATCGAGGTGATCAATGTTGATGGGACGAAGAGCGGACCTCTCTGGAGGAC 1680
Db |||||
1640 ACAGACAAGATCGAGGTGATCAATGTTGATGGGACGAAGAGCGGACCTCTCTGGAGGAC 1699
QY |||||
1681 AAGCTCCGCGACATTTTCGGGTTTCACGCTGCTGGGGGACTTCATCTACTGAGTGAATGG 1740
Db |||||
1700 AAGCTCCGCGACATTTTCGGGTTTCACGCTGCTGGGGGACTTCATCTACTGAGTGAATGG 1759
QY |||||
1741 CAGCCCGCAGCATCGAGCGGTGACAAAGTCAAGGTCAGAGCCAGCCGAGGCTCATCATTTGAC 1800
Db |||||
1760 CAGCCCGCAGCATCGAGCGGTGACAAAGTCAAGGTCAGAGCCAGCCGAGGCTCATCATTTGAC 1819
QY |||||
1801 CAGCTGCCCGACCTGATGGGGCTCAAGAGTGTGAATGTGGCCAAAGGTGCTCGGAACCAAC 1860
Db |||||
1820 CAGCTGCCCGACCTGATGGGGCTCAAGAGTGTGAATGTGGCCAAAGGTGCTCGGAACCAAC 1879
QY |||||
1861 CCGTGTGGGACAGGAAACGGGGGTGACGCCACCTGTGTGTTTTCACACCCACGCAACC 1920
Db |||||
1880 CCGTGTGGGACAGGAAACGGGGGTGACGCCACCTGTGTGTTTTCACACCCACGCAACC 1939
QY |||||
1921 CGTGTGCTGCCCATCGGCTGGAGCTGTGAGTGATGACATGACACCTGATCGTGCCT 1980
Db |||||
1940 CGTGTGCTGCCCATCGGCTGGAGCTGTGAGTGATGACATGACACCTGATCGTGCCT 1999
QY |||||
1981 GAGGCTTTCTTGGTCTTACCAGCAGAGCGCGCATCCACAGGATCTCCCTCGAGACCAAT 2040
Db |||||
2000 GAGGCTTTCTTGGTCTTACCAGCAGAGCGCGCATCCACAGGATCTCCCTCGAGACCAAT 2059
QY |||||
2041 AACACGAGCTGGCCATCCCGCTCAGCGGCGTCAAGGAGGCTCAGCCCTGGACTTTGAT 2100
Db |||||
2060 AACACGAGCTGGCCATCCCGCTCAGCGGCGTCAAGGAGGCTCAGCCCTGGACTTTGAT 2119
QY |||||
2101 GTGTCCAAACACCATCTACTGACAGACGCTCAGCCTGAAGACCATCAGCGCGGCTTC 2160
Db |||||
2120 GTGTCCAAACACCATCTACTGACAGACGCTCAGCCTGAAGACCATCAGCGCGGCTTC 2179
QY |||||
2161 ATGAACCGGAGCTGGTGGAGACAGTGTGGAGTTTGGCTTGACTACCCCGAGGGCATG 2220
Db |||||
2180 ATGAACCGGAGCTGGTGGAGACAGTGTGGAGTTTGGCTTGACTACCCCGAGGGCATG 2239
QY |||||
2221 GCCGTGAGTGGAGGCAAGACCTCTACTTGGGCGGACACTGGGACCAACAGATCGAA 2280
Db |||||
2240 GCCGTGAGTGGAGGCAAGACCTCTACTTGGGCGGACACTGGGACCAACAGATCGAA 2299
QY |||||
2281 GTGGCGGCTGGACGGGCAAGTTCGGCAAGTCTCGTGTGGAGGAGCTTGGACAAACCG 2340
Db |||||
2300 GTGGCGGCTGGACGGGCAAGTTCGGCAAGTCTCGTGTGGAGGAGCTTGGACAAACCG 2359
QY |||||
2341 AGTGTGCTGGCTGGATCCCAACGAAGGCTACATCTACTGGACCGAGTGGGGCGGCAAG 2400
Db |||||
2360 AGTGTGCTGGCTGGATCCCAACGAAGGCTACATCTACTGGACCGAGTGGGGCGGCAAG 2419
QY |||||
2401 CCGAGGATCGGCGGCTTCATGACGGGACCAACTGCATGACGCTGTGGACAAAGTG 2460
Db |||||
2420 CCGAGGATCGGCGGCTTCATGACGGGACCAACTGCATGACGCTGTGGACAAAGTG 2479
QY |||||
2461 GGCGGGGCAACGACCTCACATTTGACTACGCTGACACGAGCCCTCTACTGGACCGACCTG 2520
Db |||||
2480 GGCGGGGCAACGACCTCACATTTGACTACGCTGACACGAGCCCTCTACTGGACCGACCTG 2539
QY |||||
2521 GACACCAACATGATGAGTCTGCAACATGCTGGGTGAGGAGCGGCTGATTTGCCGAC 2580
Db |||||
2540 GACACCAACATGATGAGTCTGCAACATGCTGGGTGAGGAGCGGCTGATTTGCCGAC 2599
QY |||||
2581 GATCTCCGCGACCGCTTCCGCTGTCACGAGTACAGCGATTTATATCTACTGACAGACTGG 2640
Db |||||
2600 GATCTCCGCGACCGCTTCCGCTGTCACGAGTACAGCGATTTATATCTACTGACAGACTGG 2659
QY |||||
2641 AATCTGCACAGCATTTGAGCGGGCCGACAAAGACTAGCGCGCGGAAACCGCACCTCTATCCAG 2700

Db |||||
2660 AATCTGCACAGCATTTGAGCGGGCCGACAAAGACTAGCGCCCGGAAACCGCACCTCTATCCAG 2719
QY |||||
2701 GGCCACCTGGAGTTCGTGATGGAATCTCTGTGTTTCACTCTCTCCGCGAGGATGGCTC 2760
Db |||||
2720 GGCCACCTGGAGTTCGTGATGGAATCTCTGTGTTTCACTCTCTCCGCGAGGATGGCTC 2779
QY |||||
2761 AATGACTGTATGCAACAACCGGCGAGTGTGGGAGCTGTGCTTGGCCATCCCCGGCGGC 2820
Db |||||
2780 AATGACTGTATGCAACAACCGGCGAGTGTGGGAGCTGTGCTTGGCCATCCCCGGCGGC 2839
QY |||||
2821 CACCGCTGGGCTGGCTCACAATACACCTTGACCCAGCAGCCGCAATGTCAGCCG 2880
Db |||||
2840 CACCGCTGGGCTGGCTCACAATACACCTTGACCCAGCAGCCGCAATGTCAGCCG 2899
QY |||||
2881 CCCACCACTTCTTGTGCTTTCAGCCAGAAATCTGCCATCAGTCGGATGATCCCGAGCAG 2940
Db |||||
2900 CCCACCACTTCTTGTGCTTTCAGCCAGAAATCTGCCATCAGTCGGATGATCCCGAGCAG 2959
QY |||||
2941 CAGCAGCCCGGATCTCATCTCTGCTGCTGATGGAATGAGAACTGAAAGCCATCGAC 3000
Db |||||
2960 CAGCAGCCCGGATCTCATCTCTGCTGCTGATGGAATGAGAACTGAAAGCCATCGAC 3019
QY |||||
3001 TATGACCCACTGGACAAAGTTTCATCTACTGGGTGATGGCGCCAGAAATCAAGCGAGCC 3060
Db |||||
3020 TATGACCCACTGGACAAAGTTTCATCTACTGGGTGATGGCGCCAGAAATCAAGCGAGCC 3079
QY |||||
3061 AAGGACGACGGGACCCAGCCCTTTGTTTGTGACCTCTCTGAGCCAGGCAAAACCCAGAC 3120
Db |||||
3080 AAGGACGACGGGACCCAGCCCTTTGTTTGTGACCTCTCTGAGCCAGGCAAAACCCAGAC 3139
QY |||||
3121 AAGGACGACCCGACCTCAGCATCGAATCTACAGCCGGAACCTGTTCTGAGACGTGCGAG 3180
Db |||||
3140 AAGGACGACCCGACCTCAGCATCGAATCTACAGCCGGAACCTGTTCTGAGACGTGCGAG 3199
QY |||||
3181 GCCACCAATACCATCAACAGCTCCAGAGTGAAGGGAAGCCATGGGGGTGCTGCGT 3240
Db |||||
3200 GCCACCAATACCATCAACAGCTCCAGAGTGAAGGGAAGCCATGGGGGTGCTGCGT 3259
QY |||||
3241 GGGGACCGCGAAGCCAGGGCCATCGTGTCAACCGGAGCGAGGTTACTGTACTTC 3300
Db |||||
3260 GGGGACCGCGAAGCCAGGGCCATCGTGTCAACCGGAGCGAGGTTACTGTACTTC 3319
QY |||||
3301 ACCAACAATGAGGACCGGGGACCAAGATCGAAACCGCAGCCCTTGGAGCGGACCGAGCGC 3360
Db |||||
3320 ACCAACAATGAGGACCGGGGACCAAGATCGAAACCGCAGCCCTTGGAGCGGACCGAGCGC 3379
QY |||||
3361 GAGGTCTCTTTCACACCGGCTCATCGCCCTGTGCGCTGTGTTGGTGGACAAACACTG 3420
Db |||||
3380 GAGGTCTCTTTCACCAACCGGCTCATCGCCCTGTGCGCTGTGTTGGTGGACAAACACTG 3439
QY |||||
3421 GGCAGGCTGTTCTGGGTGGACCGGACCTGAAAGCGCATTTGAGAGCTGTGACTGTGAGG 3480
Db |||||
3440 GGCAGGCTGTTCTGGGTGGACCGGACCTGAAAGCGCATTTGAGAGCTGTGACTGTGAGG 3499
QY |||||
3481 GCCAACCGCTGACCTTGGAGGACCGCAACATCGTGCAGCTCTTGGGCTTGAACCATCTT 3540
Db |||||
3500 GCCAACCGCTGACCTTGGAGGACCGCAACATCGTGCAGCTCTTGGGCTTGAACCATCTT 3559
QY |||||
3541 GGCAGGATCTCTACTTGGATCGACCGCCAGCAGATGATTCGAGCGCTGTGAGAGAGACC 3600
Db |||||
3560 GGCAGGATCTCTACTTGGATCGACCGCCAGCAGCAGATGATTCGAGCGTGTGAGAGAGACC 3619
QY |||||
3601 ACCGGGACAGCGGATCTGATCCAGGCGGTGTGCGCCACCTCTACTTGGATCGATCGA 3660
Db |||||
3620 ACCGGGACAGCGGATCTGATCCAGGCGGTGTGCGCCACCTCTACTTGGATCGATCGA 3679
QY |||||
3661 GTGGAGGAGTTCAGCCTGGAGGAGTTCCTCAGCCCCACCATGTGCGCTGACAAATGTGGC 3720
Db |||||
3680 GTGGAGGAGTTCAGCCTGGAGGAGTTCCTCAGCCCCACCATGTGCGCTGACAAATGTGGC 3739
QY |||||
3721 TGCTCCACATCTGTATTGCGAAAGGTTGATGGGACCAACCGGTGCTCATGCCCCAGTCCAC 3780

Db 3740 TGCTCCCAATCTGTGTAATGTCAGAGGATGATGGGACACACAGGTGTCTCATGTGCCACAGTCCAC 3799
Qy 3781 CTCGTGCTCTCCGACAGACCTGTGACCTGTGAGAGCGCCACACCTGTCTCCCGGACAG 3840
Db 3800 CTCGTGCTCTCCGACAGACCTGTGACCTGTGAGAGCGCCACACCTGTCTCCCGGACAG 3859
Qy 3841 TTTGATGTGACACAGGAGAGATCGACTGTATTCCTCCGGGCTGTGGCGTGTGACGGCTTT 3900
Db 3860 TTTGATGTGACACAGGAGAGATCGACTGTATTCCTCCGGGCTGTGGCGTGTGACGGCTTT 3919
Qy 3901 CCCAGTGGATGACACAGAGCGACGAGAGGCGTCCCGGTGTGCTCCGGCGCCGAGTTC 3960
Db 3920 CCCAGTGGATGACACAGAGCGACGAGAGGCGTCCCGGTGTGCTCCGGCGCCGAGTTC 3979
Qy 3961 CCTCGCGCGGGGTGAGTGTGAGCTGTGACCTGCGCTGCGCTGCGAGCGGAGGAGACTGT 4020
Db 3980 CCTCGCGCGGGGTGAGTGTGAGCTGTGAGCTGTGCGCTGCGCTGCGAGCGGAGGAGACTGT 4039
Qy 4021 CAGGACCGCTCAGACGAGCGGACTGTGACGCCATCTGCTTGCCTGCCCAACAGTTCGGTGT 4080
Db 4040 CAGGACCGCTCAGACGAGGTGAGCTGTGACGCCATCTGCTTGCCTGCCCAACAGTTCGGTGT 4099
Qy 4081 GCGAGCGCGAGTGTCTCTCATCAAACAGAGTGGAGTCTCTTCCCGAGCTGTATCGAC 4140
Db 4100 GCGAGCGCGAGTGTCTCTCATCAAACAGAGTGGAGTCTCTTCCCGAGCTGTATCGAC 4159
Qy 4141 GGCTCCGAGAGCTCATGTGTGAATACCAAGCGCCCTCAGACGACAGCGCGGCCAC 4200
Db 4160 GGCTCCGAGAGCTCATGTGTGAATACCAAGCGCCCTCAGACGACAGCGCGGCCAC 4219
Qy 4201 AGCAGTGCATCGGGCCCGTCAITGGCATCATCTCTCTCTTCTGTCATGGGTGGTGC 4260
Db 4220 AGCAGTGCATCGGGCCCGTCAITGGCATCATCTCTCTCTCTCTGTCATGGGTGGTGC 4279
Qy 4261 TATTTTGTGTCAGCGCGTGGTGTGCCAGCGCTATGCGGGGGGCAACGGGCCCTTCCCG 4320
Db 4280 TATTTTGTGTCAGCGCGTGGTGTGCCAGCGCTATGCGGGGGGCAACGGGCCCTTCCCG 4339
Qy 4321 CACAGTATGTGACGGGACCCCGCAGCTGCCCTCAATTTTCATAGCCCGGGCGGTTC 4380
Db 4340 CACAGTATGTGACGGGACCCCGCAGCTGCCCTCAATTTTCATAGCCCGGGCGGTTC 4399
Qy 4381 CAGCATGGCCCTTTCACAGGCGATCGCATGCGGAAAGTCCATGATGAGCTCCGTGAGCGTG 4440
Db 4400 CAGCATGGCCCTTTCACAGGCGATCGCATGCGGAAAGTCCATGATGAGCTCCGTGAGCGTG 4459
Qy 4441 ATGGGGGGCGGGGGGGTGCCTCTGTACACCGGAACACAGTCAAGGGGGCTCGTCC 4500
Db 4460 ATGGGGGGCGGGGGGGTGCCTCTGTACACCGGAACACAGTCAAGGGGGCTCGTCC 4519
Qy 4501 AGCAGCTCGTCCAGCAGGAGCGCAGCTGTACCGCGGATCCTGAAACCGCGCCCTCC 4560
Db 4520 AGCAGCTCGTCCAGCAGGAGCGCAGCTGTACCGCGGATCCTGAAACCGCGCCCTCC 4579
Qy 4561 CCGGCCACGGAACCCCTCCCTGTACAAATGACATGTTCTACTCTTCAAAATTCGGGCC 4620
Db 4580 CCGGCCACGGAACCCCTCCCTGTACAAATGACATGTTCTACTCTTCAAAATTCGGGCC 4639
Qy 4621 ACTCGGAGACCGTACAGGCCCTCATCATTCAGGAATGGGGCCCCCGAGACGCCCTGC 4680
Db 4640 ACTCGGAGACCGTACAGGCCCTCATCATTCAGGAATGGGGCCCCCGAGACGCCCTGC 4699
Qy 4681 AGCACCGAGTGTGTGACAGGACTACAGCGCCAGCGCTGGAAGGCCAGCAAGTACTAC 4740
Db 4700 AGCACCGAGTGTGTGACAGGACTACAGCGCCAGCGCTGGAAGGCCAGCAAGTACTAC 4759
Qy 4741 CTGGATTTGAATCGGACTCAGACCCCTTATCACCCCGCCACCGCCCGCCACAGCCAGTAC 4800
Db 4760 CTGGATTTGAATCGGACTCAGACCCCTTATCACCCCGCCACCGCCCGCCACAGCCAGTAC 4819
Qy 4801 CTGTGGCGGAGGACAGCTGCCCGCCCTGCGCCCGCCACCGGAGGAGGACTTCCATCTC 4860
Db 4820 CTGTGGCGGAGGACAGCTGCCCGCCCTGCGCCCGCCACCGGAGGAGGACTTCCATCTC 4879

Qy 4861 TTCCCGCCCTCCGTCCCTCCCTCGACCGGACTCATCTGACCTCGGCGGGCCACTCTCGC 4920
Db 4880 TTCCCGCCCTCCGTCCCTCCCTCGACCGGACTCATCTGACCTCGGCGGGCCACTCTCGC 4939
Qy 4921 TTCTCTGTGCCCCCTGTAATAATAGTTTTTAAATATGAACAAAGAAAAAATATATTTATGAT 4980
Db 4940 TTCTCTGTGCCCCCTGTAATAATAGTTTTTAAATATGAACAAAGAAAAAATATATTTATGAT 4999
Qy 4981 TTAATAAATAAATAATAATTTGGATTTTAAAAACATGAGAAATGTGAATGTGATGGGTG 5040
Db 5000 TTAATAAATAAATAATAATTTGGATTTTAAAAACATGAGAAATGTGAATGTGATGGGTG 5059
Qy 5041 GGCAGGCTGGGAGAACTTTTGTGA 5063
Db 5060 GGCAGGCTGGGAGAACTTTTGTGA 5082

Search completed: February 18, 2005, 14:15:37
Job time : 2519 secs

THIS PAGE BLANK (USPTO)

Qy	301	GAGAGGCCATCAAGCAGACCTTACTTGAAACGAGACGGGGGCGCCGTGCAGAACTGTGTC	360
Db	320	GAGAGGCCATCAAGCAGACCTTACTTGAAACGAGACGGGGGCGCCGTGCAGAACTGTGTC	379
Qy	361	ATCTCCGGCCTGGTCTCTCCCGACGGGCTCGCCTCGCACTGGGTGGGCAAGACTGTAC	420
Db	380	ATCTCCGGCCTGGTCTCTCCCGACGGGCTCGCCTCGCACTGGGTGGGCAAGACTGTAC	439
Qy	421	TGGACGGACTCAGAGACAACCCGATCGAGGTGGCCAACTCAATAGGCACATCCCGGAAG	480
Db	440	TGGACGGACTCAGAGACAACCCGATCGAGGTGGCCAACTCAATAGGCACATCCCGGAAG	499
Qy	481	GTGCTCTTCGGCAGGACCTTGACAGCCTAGGGCCATCGCCTTTGGACCCCGCTCAGCGG	540
Db	500	GTGCTCTTCGGCAGGACCTTGACAGCCTAGGGCCATCGCCTTTGGACCCCGCTCAGCGG	559
Qy	541	TACATGTACTGGACAGACTGGGGTGAAGCGCCCGGATTTAGCGGGCAGGGATGGATGGC	600
Db	560	TACATGTACTGGACAGACTGGGGTGAAGCGCCCGGATTTAGCGGGCAGGGATGGATGGC	619
Qy	601	AGCACCGGAAGATCATTTGTGGACTCGGACATTTACTGGCCCAATGACATGCATCGAC	660
Db	620	AGCACCGGAAGATCATTTGTGGACTCGGACATTTACTGGCCCAATGACATGCATCGAC	679
Qy	661	CTGAGGAGCAGAAAGCTTACTCTGGGCTGACGCCAAGCTTCAGTTTCATCCACCGTGCCAA	720
Db	680	CTGAGGAGCAGAAAGCTTACTCTGGGCTGACGCCAAGCTTCAGTTTCATCCACCGTGCCAA	739
Qy	721	CTGACCGGCTGTTCCGGCAGAAAGTGGTGGAGGGCAGCTTGACGCAACCCCTTCGCCCTG	780
Db	740	CTGACCGGCTGTTCCGGCAGAAAGTGGTGGAGGGCAGCTTGACGCAACCCCTTCGCCCTG	799
Qy	781	ACGCTCTCCGGGGACCTCTGTACTGACAGACTGGCAGACCCGCTCCATCCATGCTCTGC	840
Db	800	ACGCTCTCCGGGGACCTCTGTACTGACAGACTGGCAGACCCGCTCCATCCATGCTCTGC	859
Qy	841	AACAAGCGCACTGGGGGGAAGAGGAGATCTGTAGTGCCCTTCTACTCACCCATGGAC	900
Db	860	AACAAGCGCACTGGGGGGAAGAGGAGATCTGTAGTGCCCTTCTACTCACCCATGGAC	919
Qy	901	ATCCAGGTGTCAGCAGGACGGCAGCCTTTCTTCCACACTCGCTGTGAGGAGCAAT	960
Db	920	ATCCAGGTGTCAGCAGGACGGCAGCCTTTCTTCCACACTCGCTGTGAGGAGCAAT	979
Qy	961	GGCGGCTGCTCCCACTGTGCTGTCTGCCAAGCGAGCCTTTCTACATGCGCCTGC	1020
Db	980	GGCGGCTGCTCCCACTGTGCTGTCTGCCAAGCGAGCCTTTCTACATGCGCCTGC	1039
Qy	1021	CCCAAGGGTGTGAGTGTGAGGACAAAGCGCAGGACGTTGAGGACGAGCGGAGGGTG	1080
Db	1040	CCCAAGGGTGTGAGTGTGAGGACAAAGCGCAGGACGTTGAGGACGAGCGGAGGGTG	1099
Qy	1081	CTGCTGTGGCCGGCGGACGGACCTACGAGGATCTCGTTGGACACGCGGACTTCACC	1140
Db	1100	CTGCTGTGGCCGGCGGACGGACCTACGAGGATCTCGTTGGACACGCGGACTTCACC	1159
Qy	1141	GACATCGTGTGAGGTGAGCAGATCCGGCAGCCCATTCGCACTACGACTACGACCGGCTA	1200
Db	1160	GACATCGTGTGAGGTGAGCAGATCCGGCAGCCCATTCGCACTACGACTACGACCGGCTA	1219
Qy	1201	GAGGGCTATGTCTACTGGAAGATGAGAGGTGGGGCCATCCGACGGGGGTACTTGGAC	1260
Db	1220	GAGGGCTATGTCTACTGGAAGATGAGAGGTGGGGCCATCCGACGGGGGTACTTGGAC	1279
Qy	1261	GGGTCTGGGGCGCAGACGCTGGTCAACACCGAGATCAACGACCCCGATGGCATCGCGGTC	1320
Db	1280	GGGTCTGGGGCGCAGACGCTGGTCAACACCGAGATCAACGACCCCGATGGCATCGCGGTC	1339
Qy	1321	GACTGGGTGGCCCGAAACCTTACTTGACACGACACCGGGCAGGACCGGATCGAGGTGACG	1380
Db	1340	GACTGGGTGGCCCGAAACCTTACTTGACACGACACCGGGCAGGACCGGATCGAGGTGACG	1399

Qy	1381	CGCCTCAACGGCACCTTCCGCAAGATCCTGGTGTGCGAGGACCTTGGACGACGACCCCGAGCC	1444
Db	1400	CGCCTCAACGGCACCTTCCGCAAGATCCTGGTGTGCGAGGACCTTGGACGACGACCCCGAGCC	1459
Qy	1441	ATCGCACTGCACCCCGTGATGGCGCTCATGTACTGGACAGACTGGGAGAGAAACCCCTAAA	1500
Db	1460	ATCGCACTGCACCCCGTGATGGCGCTCATGTACTGGACAGACTGGGAGAGAACCCCTAAA	1519
Qy	1501	ATCGAGTGTGCCAACTTTGGATGGGCAAGGCGCGTGTGCTGTCAATGCTCCTCCCTCGGG	1560
Db	1520	ATCGAGTGTGCCAACTTTGGATGGGCAAGGCGCGTGTGCTGTCAATGCTCCTCCCTCGGG	1579
Qy	1561	TGGCCCAACGGCCTTGGCCCTTGGACCTTGCAGAGGGGAAGCTCTACTTGGGAGAGCGCCAAG	1620
Db	1580	TGGCCCAACGGCCTTGGCCCTTGGACCTTGCAGAGGGGAAGCTCTACTTGGGAGAGCGCCAAG	1639
Qy	1621	ACAGACAAGATCGAGGTGATCAATGTTGTATGGGACGAAGAGCGGAGCCCTCTCTGGAGGAC	1680
Db	1640	ACAGACAAGATCGAGGTGATCAATGTTGTATGGGACGAAGAGCGGAGCCCTCTCTGGAGGAC	1699
Qy	1681	AAGCTCCGCGACATTTTCGGGTTTCACGTCTCTGGGGACTTTCACTACTTGGACTGACTGG	1740
Db	1700	AAGCTCCGCGACATTTTCGGGTTTCACGTCTCTGGGGACTTTCACTACTTGGACTGACTGG	1759
Qy	1741	CAGCGCGCAGCATCGAGCGGGTGCAACAAGGTCAAAGCCAGCCGCGGACGTCATCATTTGAC	1800
Db	1760	CAGCGCGCAGCATCGAGCGGGTGCAACAAGGTCAAAGCCAGCCGCGGACGTCATCATTTGAC	1819
Qy	1801	CAGCTGCCCGACCTGATGGGGCTCAAAGCTGTGAATGTGGCCAAAGTCTGTGGAAACCAAC	1860
Db	1820	CAGCTGCCCGACCTGATGGGGCTCAAAGCTGTGAATGTGGCCAAAGTCTGTGGAAACCAAC	1879
Qy	1861	CCGTGTGCGGACAGGAACGGGGGGTGAGCCACTGTGCTTTCTTCAACCCACGCAAC	1920
Db	1880	CCGTGTGCGGACAGGAACGGGGGGTGAGCCACTGTGCTTTCTTCAACCCACGCAAC	1939
Qy	1921	CGGTGTGGCTGCCCATCGGCCCTTGGAGCTGCTGAGTGAACATGAAGACCTGCATCGTGCCT	1980
Db	1940	CGGTGTGGCTGCCCATCGGCCCTTGGAGCTGCTGAGTGAACATGAAGACCTGCATCGTGCCT	1999
Qy	1981	GAGGCTTTCTTGCTCTTCAACGACGAGCGCCATCCACAGGATCTCCCTCGAGAGCCAAT	2040
Db	2000	GAGGCTTTCTTGCTCTTCAACGACGAGCGCCATCCACAGGATCTCCCTCGAGAGCCAAT	2059
Qy	2041	AACAAACGACGTGCCCATCCCGCTCACGGGCGTCAAGGAGGCTTCAGCCCTTGGACTTTGAT	2100
Db	2060	AACAAACGACGTGCCCATCCCGCTCACGGGCGTCAAGGAGGCTTCAGCCCTTGGACTTTGAT	2119
Qy	2101	GTGTCCAAACACACATCTACTTGGACAGAGCTCAGCTGAGCAATGAGCAATCAGCCGCTTC	2160
Db	2120	GTGTCCAAACACACATCTACTTGGACAGAGCTCAGCTGAGCAATGAGCAATCAGCCGCTTC	2179
Qy	2161	ATGAAACGGGAGCTCGGTGGAGACGTGTGGAGTTTGGCTTCACTACCCGAGGGCATG	2220
Db	2180	ATGAAACGGGAGCTCGGTGGAGACGTGTGGAGTTTGGCTTCACTACCCGAGGGCATG	2239
Qy	2221	GCCGTTGACTGATGGGCAAGAACTCTACTGGCCGCACTCTGGGACCAACAGAAATCGAA	2280
Db	2240	GCCGTTGACTGATGGGCAAGAACTCTACTGGCCGCACTCTGGGACCAACAGAAATCGAA	2299
Qy	2281	GTGCGCGGCTGACGGGCGATTTCCGGCAAGTCTCGTGTGGAGGACTTTGGACAAACCCG	2340
Db	2300	GTGCGCGGCTGACGGGCGATTTCCGGCAAGTCTCGTGTGGAGGACTTTGGACAAACCCG	2359
Qy	2341	AGGTTCGTGGCCCTGGATCCCAACAAAGGGCTACATCTACTGGACCGAGTGGGCGGCAAG	2400
Db	2360	AGGTTCGTGGCCCTGGATCCCAACAAAGGGCTACATCTACTGGACCGAGTGGGCGGCAAG	2419
Qy	2401	CCGAGGATCGTGGGGCTTCATGGACGGGACCAACTGCATGACGCTGTGGACAGAGTG	2460
Db	2420	CCGAGGATCGTGGGGCTTCATGGACGGGACCAACTGCATGACGCTGTGGACAGAGTG	2479
Qy	2461	GGCGGGGCAACGACCTCACCATTTGACTACGCTGACGACGCGCTCTACTTGGACCGGACCTG	2520


```
QY 4691 AGCACCGACGTGTGTGACAGCGACTACAGCGCAGCCCTCGAAGGCCAGCAAGTACTAC 4740
DB 4700 AGCACCGACGTGTGTGACAGCGACTACAGCGCAGCCCTCGAAGGCCAGCAAGTACTAC 4759
QY 4741 CTGGATTTGAATCGGACTCAGACCCCTATCCACCCCCACCCAGCCAGCCAGCCAGTAC 4800
DB 4760 CTGGATTTGAATCGGACTCAGACCCCTATCCACCCCCACCCAGCCAGCCAGTACTTCCATCTC 4819
QY 4801 CTGTGCGCGGAGGACAGCTGCGCCCTCGCCCGCCACCGAGAGGAGTACTTCCATCTC 4860
DB 4820 CTGTGCGCGGAGGACAGCTGCGCCCTCGCCCGCCACCGAGAGGAGTACTTCCATCTC 4879
QY 4861 TTCGCCCCCTCCGTCCCTCGACGACTCATCTGACCTCGCCCGGCGACTCTGGC 4920
DB 4880 TTCGCCCCCTCCGTCCCTCGACGACTCATCTGACCTCGCCCGGCGACTCTGGC 4939
QY 4921 TTCTCTGCGCCCTGTAATAGTTTAAATATGAACAAAGAAAAATATATTTATGAT 4980
DB 4940 TTCTCTGCGCCCTGTAATAGTTTAAATATGAACAAAGAAAAATATATTTATGAT 4999
QY 4981 TTAATAAATAAATAAATGGAATTTAAACATGAGAAATGTGAATGTGATGGGTG 5040
DB 5000 TTAATAAATAAATAAATGGAATTTAAACATGAGAAATGTGAATGTGATGGGTG 5059
QY 5041 GCGAGGCTGGGAGACTTTGTA 5063
DB 5060 GCGAGGCTGGGAGACTTTGTA 5082
```

RESULT 2

US-09-543-771B-1

; Sequence 1, Application US/09543771B

; Patent No. 6780609

; GENERAL INFORMATION:

; APPLICANT: Carulli, John P.

; APPLICANT: Little, Randall D.

; APPLICANT: Recker, Robert L.

; APPLICANT: Johnson, Mark L.

; TITLE OF INVENTION: High bone mass gene of 11q13.3

; FILE REFERENCE: 032796-014

; CURRENT APPLICATION NUMBER: US/09/543,771B

; CURRENT FILING DATE: 2000-04-05

; PRIOR APPLICATION NUMBER: US 09/229,319

; PRIOR FILING DATE: 1999-01-13

; PRIOR APPLICATION NUMBER: US 60/071,449

; PRIOR FILING DATE: 1998-01-13

; PRIOR APPLICATION NUMBER: US 60/105,511

; PRIOR FILING DATE: 1998-10-23

; NUMBER OF SEQ ID NOS: 641

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1

; LENGTH: 5120

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-543-771B-1

Query Match

Best Local Similarity 99.9%; Score 5058.2; DB 4; Length 5120;

Matches 5060; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
QY 1 GCCATGGAGCCGAGTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCA 60
DB 20 GCCATGGAGCCGAGTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCA 79
QY 61 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
DB 80 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 139
QY 121 GCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
DB 140 GCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 199
```

```
QY 181 CCGCTGGTGGACCGCGCGGAGTCAAGCTGGAGTCCACCATCGTGTGTCAGCGGCGCTGAG 240
DB 200 CCGCTGGTGGACCGCGCGGAGTCAAGCTGGAGTCCACCATCGTGTGTCAGCGGCGCTGAG 259
QY 241 GATCGCGCGCAGTGGAGTCTTCCAGTGTTCGAAGGAGCGGTGTACTGTGACAGAGCGTGAGC 300
DB 260 GATCGCGCGCAGTGGAGTCTTCCAGTGTTCGAAGGAGCGGTGTACTGTGACAGAGCGTGAGC 319
QY 301 GAGGAGGCCATCAAGACAGACCTACTGAAACAGACGGGGGCCCGGTGACAGAAAGTGTGC 360
DB 320 GAGGAGGCCATCAAGACAGACCTACTGAAACAGACGGGGGCCCGGTGACAGAAAGTGTGC 379
QY 361 ATCTCCGGCCTGGTCTCTCCCGACGGGCTGCGCTGCGACTGGGTGGGCAAGAGCTGAC 420
DB 380 ATCTCCGGCCTGGTCTCTCCCGACGGGCTGCGCTGCGACTGGGTGGGCAAGAGCTGAC 439
QY 421 TGGACGGACTCAGAGACCAACCGCATCGAGTGGCCCAACCTCAATGGCACATCCCGGAAG 480
DB 440 TGGACGGACTCAGAGACCAACCGCATCGAGTGGCCCAACCTCAATGGCACATCCCGGAAG 499
QY 481 GTGCTCTTCTGGCAGGACCTTTGACAGCCTAGAGGCCATGCGCTTTGGACCCCGCTCACGGG 540
DB 500 GTGCTCTTCTGGCAGGACCTTTGACAGCCTAGAGGCCATGCGCTTTGGACCCCGCTCACGGG 559
QY 541 TACATGTACTGACAGACTGGGTGAGACGCCCGCGATTGAGCGGGCAGGGATGATGCGC 600
DB 560 TACATGTACTGACAGACTGGGTGAGACGCCCGCGATTGAGCGGGCAGGGATGATGCGC 619
QY 601 AGCACCCGGGAAGATCATTTGTGGAATCTCGGACATTTTACTGGCCCAATGGAATGACCATCGAC 660
DB 620 AGCACCCGGGAAGATCATTTGTGGAATCTCGGACATTTTACTGGCCCAATGGAATGACCATCGAC 679
QY 661 CTGAGGAGCAGAAAGCTCTACTGGGTGAGACGCCAAGCTCAGCTTCATCCACCGTGCCCAAC 720
DB 680 CTGAGGAGCAGAAAGCTCTACTGGGTGAGACGCCAAGCTCAGCTTCATCCACCGTGCCCAAC 739
QY 721 CTGAGCGGCTGTTCCGGCAGAGGTGTGGAGGCGAGCCTGAGCGACCCCTTCGCCCTG 780
DB 740 CTGAGCGGCTGTTCCGGCAGAGGTGTGGAGGCGAGCCTGAGCGACCCCTTCGCCCTG 799
QY 781 ACGCTCTCCGGGACACTCTGTACTGGACAGACTGGCAGACCCCGCTCCATCATGCGCTGC 840
DB 800 ACGCTCTCCGGGACACTCTGTACTGGACAGACTGGCAGACCCCGCTCCATCATGCGCTGC 859
QY 841 AACAGCGCATGCGGGGAGAGAGAGAGATCTGAGTGCCCTCTACTCAACCATGAC 900
DB 860 AACAGCGCATGCGGGGAGAGAGAGAGATCTGAGTGCCCTCTACTCAACCATGAC 919
QY 901 ATCCAGGTGTGAGCCAGGAGCGGACGCTTTTCCACACTCGCTGTGAGGAGACAAAT 960
DB 920 ATCCAGGTGTGAGCCAGGAGCGGACGCTTTTCCACACTCGCTGTGAGGAGACAAAT 979
QY 961 GCGGCTGTCTCCACCTGTGCTGTCTGTCTCCCAAGCGAGCGCTTTTACACATGCGCTGC 1020
DB 980 GCGGCTGTCTCCACCTGTGCTGTCTGTCTCCCAAGCGAGCGCTTTTACACATGCGCTGC 1039
QY 1021 CCCACGGGTGTGAGCTGACAGCAACCGGACAGAGCTGTAGGAGGAGCGCGAGAGGTG 1080
DB 1040 CCCACGGGTGTGAGCTGACAGCAACCGGACAGAGCTGTAGGAGGAGCGCGAGAGGTG 1099
QY 1081 CTGCTGTGCGCCCGCGGACGACCTACGGAGGATCTCGCTGGACACCGCGGACTTCACC 1140
DB 1100 CTGCTGTGCGCCCGCGGACGACCTACGGAGGATCTCGCTGGACACCGCGGACTTCACC 1159
QY 1141 GACATCGTGTGACAGGTGAGCAGCATCCGGACGCCATTCATCGACTACGACCCGCTA 1200
DB 1160 GACATCGTGTGACAGGTGAGCAGCATCCGGACGCCATTCATCGACTACGACCCGCTA 1219
QY 1201 GAGGCTATGTCTACTGACAGATGACAGGTGGGGGCGCATCCCGAGGGCGCTACTGAGC 1260
DB 1220 GAGGCTATGTCTACTGACAGATGACAGGTGGGGGCGCATCCCGAGGGCGCTACTGAGC 1279
QY 1261 GGGTCTGGGGCGCAGACGCTGTGTTCAACACCGAGATCAACGACCCCGATGGCATGCGCGTGC 1320
```


QY	1081	CTGCTGCTGGCCCGCGGACGACCTACGAGGATCTCGTGGACACGCGGACTTACCC	1140
Db	1100	CTGCTGCTGGCCCGCGGACGACCTACGAGGATCTCGTGGACACGCGGACTTACCC	1159
QY	1141	GACATCGTGTGAGTGGACGACATCCGGACGCAATTCGATCGATACGACCCGCTA	1200
Db	1160	GACATCGTGTGAGTGGACGACATCCGGACGCAATTCGATCGATACGACCCGCTA	1219
QY	1201	GAGGCTATGCTTACTGGACAGATGACGAGTGGCGGCAATCCGACGGCGTACTGGAC	1260
Db	1220	GAGGCTATGCTTACTGGACAGATGACGAGTGGCGGCAATCCGACGGCGTACTGGAC	1279
QY	1261	GGTCTGGGGCGCAGACCTGGTCAACACCGAGATCAAGACCCCGATGGCATCGCGTC	1320
Db	1280	GGTCTGGGGCGCAGACCTGGTCAACACCGAGATCAAGACCCCGATGGCATCGCGTC	1339
QY	1321	GACTGGGTGGCCCGGAAACCTTACTGGACCGACACGGGACCGCATCGAGTGAAG	1380
Db	1340	GACTGGGTGGCCCGGAAACCTTACTGGACCGACACGGGACCGCATCGAGTGAAG	1399
QY	1381	GGCTTCAACGGGACCTCCGCAAGATCTGTGTGCGAGGACCTGGACGAGCCCGAGCC	1440
Db	1400	GGCTTCAACGGGACCTCCGCAAGATCTGTGTGCGAGGACCTGGACGAGCCCGAGCC	1459
QY	1441	ATCGACCTGACCCCGTGCATGGGCTCATGTACTGGACAGACTGGGGAGAGAACCTTAA	1500
Db	1460	ATCGACCTGACCCCGTGCATGGGCTCATGTACTGGACAGACTGGGGAGAGAACCTTAA	1519
QY	1501	ATCGAGTGTGCAACTTGGATGGGCGAGGCGGCTGTGCTCAATGCTCCTCGGG	1560
Db	1520	ATCGAGTGTGCAACTTGGATGGGCGAGGCGGCTGTGCTCAATGCTCCTCGGG	1579
QY	1561	TGGCCCAACGGGCTGGCCCTGACCTGACGAGGGGAGGCTTACTTGGGAGAGCCCAAG	1620
Db	1580	TGGCCCAACGGGCTGGCCCTGACCTGACGAGGGGAGGCTTACTTGGGAGAGCCCAAG	1639
QY	1621	ACAGACAGATCGAGTGCATCAATGTGATGGGACGAGGCGGACCTCTCGAGGAC	1680
Db	1640	ACAGACAGATCGAGTGCATCAATGTGATGGGACGAGGCGGACCTCTCGAGGAC	1699
QY	1681	AAGCTCCCGCACATTTTCGGGTTTCAACGCTGTGGGGGACTTTCATCTACTGGACTGCTG	1740
Db	1700	AAGCTCCCGCACATTTTCGGGTTTCAACGCTGTGGGGGACTTTCATCTACTGGACTGCTG	1759
QY	1741	CAGCGCCGACGATCGAGCGGGTGCAAAAGTCAAGGCGAGCGGAGCGTCAATTTGAC	1800
Db	1760	CAGCGCCGACGATCGAGCGGGTGCAAAAGTCAAGGCGAGCGGAGCGTCAATTTGAC	1819
QY	1801	CAGCTGCCGACCTGATGGGCTCAAAAGTGTGATGTGGCCAAAGTGTGGAGCCAAAC	1860
Db	1820	CAGCTGCCGACCTGATGGGCTCAAAAGTGTGATGTGGCCAAAGTGTGGAGCCAAAC	1879
QY	1861	CCGTGTGGGACAGGAAACGGGGGGTGACACCACTGTGCTTTCACACCCCAAGCAACC	1920
Db	1880	CCGTGTGGGACAGGAAACGGGGGGTGACACCACTGTGCTTTCACACCCCAAGCAACC	1939
QY	1921	CGTGTGGCTGCCCATCGGCTGTGAGTGTGATGATGATGATGATGATGATGATGATGATG	1980
Db	1940	CGTGTGGCTGCCCATCGGCTGTGAGTGTGATGATGATGATGATGATGATGATGATGATG	1999
QY	1981	GAGGCTTTCTTGTGTTTCCACGAGAGGCGCCCATCCACAGGATCTCCCTCGAGACCAAT	2040
Db	2000	GAGGCTTTCTTGTGTTTCCACGAGAGGCGCCCATCCACAGGATCTCCCTCGAGACCAAT	2059
QY	2041	AACAAACGAGTGGCCATCCCGCTCAACGGGCGTCAAGGAGGCTCAGCCCTGGACTTTGAT	2100
Db	2060	AACAAACGAGTGGCCATCCCGCTCAACGGGCGTCAAGGAGGCTCAGCCCTGGACTTTGAT	2119
QY	2101	GTGTCCAAACCAACATCTACTGACAGAGCTGACGCTGAGAGCAGATCAGGCGGCTTC	2160
Db	2120	GTGTCCAAACCAACATCTACTGACAGAGCTGACGCTGAGAGCAGATCAGGCGGCTTC	2179

QY	2161	ATGAACGGGAGCTCGTGGAGCAGTGGTGGAGTTTGGCTTGACTACCCCGAGGGCATG	2220
Db	2180	ATGAACGGGAGCTCGTGGAGCAGTGGTGGAGTTTGGCTTGACTACCCCGAGGGCATG	2239
QY	2221	GCCGTGTGACTGGATGGGCAAGAACTTCTACTGGGCGGACACTGGGACCAACAGAACTGAA	2280
Db	2240	GCCGTGTGACTGGATGGGCAAGAACTTCTACTGGGCGGACACTGGGACCAACAGAACTGAA	2299
QY	2281	GTGGCGGCTGTGACGCGGACGTTCCGGCAAGTCTCTGTGTGGAGGAGCTTGGACAAACCG	2340
Db	2300	GTGGCGGCTGTGACGCGGACGTTCCGGCAAGTCTCTGTGTGGAGGAGCTTGGACAAACCG	2359
QY	2341	AGGTGCTGTCGCTGGATCCCAACAGGCTTACATCTACTGACCGAGTGGGGGGCAAG	2400
Db	2360	AGGTGCTGTCGCTGGATCCCAACAGGCTTACATCTACTGACCGAGTGGGGGGCAAG	2419
QY	2401	CCGAGGATCTGTGCGGGCTTTCATGACGCGGACCAACTGATGACGCTGGTGACAAAGGTG	2460
Db	2420	CCGAGGATCTGTGCGGGCTTTCATGACGCGGACCAACTGATGACGCTGGTGACAAAGGTG	2479
QY	2461	GGCGGGGCAACGACCTTCAACATTTGACTTACGCTGACGAGCGCTTACTGACCGACCTG	2520
Db	2480	GGCGGGGCAACGACCTTCAACATTTGACTTACGCTGACGAGCGCTTACTGACCGACCTG	2539
QY	2521	GACACCAACATGATCGAGTCTGCTCAACATGCTGGGTCAAGGAGCGGCTGATTTGCCGAC	2580
Db	2540	GACACCAACATGATCGAGTCTGCTCAACATGCTGGGTCAAGGAGCGGCTGATTTGCCGAC	2599
QY	2581	GATCTCCCGCACCCGTTTGGTCTGACGAGTACAGCGATTATATCTACTGACGAGACTGG	2640
Db	2600	GATCTCCCGCACCCGTTTGGTCTGACGAGTACAGCGATTATATCTACTGACGAGACTGG	2659
QY	2641	AATCTGACAGCATTGAGCGGGCGGACAAAGCTAGCGGCGGAAACCGACCCCTCATCCAG	2700
Db	2660	AATCTGACAGCATTGAGCGGGCGGACAAAGCTAGCGGCGGAAACCGACCCCTCATCCAG	2719
QY	2701	GGCCACCTTGGACTTGTGATGAGCATCTCTGTGTTCACCTCTCCCGCAGGATGGCTC	2760
Db	2720	GGCCACCTTGGACTTGTGATGAGCATCTCTGTGTTCACCTCTCCCGCAGGATGGCTC	2779
QY	2761	AATGACTGTATGACAAACAGGCGAGTGTGGGAGCTGTGCTTGGCCATCCCGGCGGC	2820
Db	2780	AATGACTGTATGACAAACAGGCGAGTGTGGGAGCTGTGCTTGGCCATCCCGGCGGC	2839
QY	2821	CACGCTGCGGCTGCGCTTACACTACCCCTGACGCCAGCGGACCGGAACTGACGCGCG	2880
Db	2840	CACGCTGCGGCTGCGCTTACACTACCCCTGACGCCAGCGGACCGGAACTGACGCGCG	2899
QY	2881	CCACACCTTCTTGTGTTTTCAGCCAGAAATCTGCCATCAGTCCGATGATCCCGGACGAC	2940
Db	2900	CCACACCTTCTTGTGTTTTCAGCCAGAAATCTGCCATCAGTCCGATGATCCCGGACGAC	2959
QY	2941	CAGCACAGCCGGAATCTCATCTCTCCCTTGCATGAGTGTGAGGAACTGTAAGCCATCGAC	3000
Db	2960	CAGCACAGCCGGAATCTCATCTCTCCCTTGCATGAGTGTGAGGAACTGTAAGCCATCGAC	3019
QY	3001	TATGACCCACTGGAAGATTTTCTACTTGGTGTGATGGGCGGAGAACTCAAGGAGCC	3060
Db	3020	TATGACCCACTGGAAGATTTTCTACTTGGTGTGATGGGCGGAGAACTCAAGGAGCC	3079
QY	3061	AAGGACGAGGAGCCAGCCCTTTTGTGACTTCTGTAGCCAAAGGCGGAGGAGCCAGAC	3120
Db	3080	AAGGACGAGGAGCCAGCCCTTTTGTGACTTCTGTAGCCAAAGGCGGAGGAGGAGCCAGAC	3139
QY	3121	AGGACGAGCCCAAGCTTCTGAGCATCGATCTTACAGCGGAGCACTGTTTCTGAGCGTGCAG	3180
Db	3140	AGGACGAGCCCAAGCTTCTGAGCATCGATCTTACAGCGGAGCACTGTTTCTGAGCGTGCAG	3199
QY	3181	GCCACCAATACCACTCAAGCTTCCAGGCTGAGCGGAGGAGGAGGAGGAGGAGGAGGAGG	3240
Db	3200	GCCACCAATACCACTCAAGCTTCCAGGCTGAGCGGAGGAGGAGGAGGAGGAGGAGGAGG	3259
QY	3241	GGGACCGCGACAAAGCCAGGGGCTTCTGCTCAACGCGGAGGAGGAGGAGGAGGAGGAGG	3300

Db 3260 GGGGACCGGCAAGCCAGCCAGGCGCATCGTCTCAACGCGGAGGAGGCTACTGTTACTTTC 3319
QY 3301 ACCACATGCGAGGCGGCGGAGCCAGATCGAAGCGGAGCGGCGGAGCGGCGGAGCGGCGG 3360
Db 3320 ACCACATGCGAGGCGGCGGAGCCAGATCGAAGCGGAGCGGCGGAGCGGCGGAGCGGCGG 3379
QY 3361 GAGGTCTCTTTCACACACCGGCGCTCATCGCGCCCTGTGCGCCCTGTGCGGCGGCGGAGCGGCGG 3420
Db 3380 GAGGTCTCTTTCACACACCGGCGCTCATCGCGCCCTGTGCGCCCTGTGCGGCGGCGGAGCGGCGG 3439
QY 3421 GCGAAGCTGTTCTGGGTGGAGCGCGGACCTGGAAGCGCATTTGAGAGCTGTGACTGTGCGG 3480
Db 3440 GCGAAGCTGTTCTGGGTGGAGCGCGGACCTGGAAGCGCATTTGAGAGCTGTGACTGTGCGG 3499
QY 3481 GCCAACCGCCTGACCGCTGGAGGCGCCAACTGCTGAGCGCTCTGGGCGCTGACCATCTT 3540
Db 3500 GCCAACCGCCTGACCGCTGGAGGCGCCAACTGCTGAGCGCTCTGGGCGCTGACCATCTT 3559
QY 3541 GCGAAGCATCTTACTGATCGACCGCGGAGCGGAGATGATCGAGCGGTGTGGAGAGACC 3600
Db 3560 GCGAAGCATCTTACTGATCGACCGCGGAGCGGAGATGATCGAGCGGTGTGGAGAGACC 3619
QY 3601 ACCGGGGAAGAAGCGGACTCGCATCCAGGCGGCTGTGCGCCACCTCACTGGCATCCATGCA 3660
Db 3620 ACCGGGGAAGAAGCGGACTCGCATCCAGGCGGCTGTGCGCCACCTCACTGGCATCCATGCA 3679
QY 3661 GTGGAGGAGTCAAGCGGAGGAGTCTCAGCGCCACCATGTGCGCGGTGACAAATGGTGGC 3720
Db 3680 GTGGAGGAGTCAAGCGGAGGAGTCTCAGCGCCACCATGTGCGCGGTGACAAATGGTGGC 3739
QY 3721 TGCTCCCAACATCTGATTGCGCAAGGAGTATGGGACACCAAGCGGTGTCTATGCGCGGAGTCCAC 3780
Db 3740 TGCTCCCAACATCTGATTGCGCAAGGAGTATGGGACACCAAGCGGTGTCTATGCGCGGAGTCCAC 3799
QY 3781 CTCGTGCTCTGCGAAGCCTGCTGACCTGTGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3840
Db 3800 CTCGTGCTCTGCGAAGCCTGCTGACCTGTGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3859
QY 3841 TTTTCATGTGCGACAGGAGGAGTCAAGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3900
Db 3860 TTTTCATGTGCGACAGGAGGAGTCAAGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3919
QY 3901 CCCGAGTCCGATGACAGAGCGGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3960
Db 3920 CCCGAGTCCGATGACAGAGCGGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3979
QY 3961 CCCTGCGCGCGGCGGCGGAGTGTGAGCCTGCGCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGG 4020
Db 3980 CCCTGCGCGCGGCGGCGGAGTGTGAGCCTGCGCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGG 4039
QY 4021 CAGGACCGCTCAGAGCGGCGGAGTGTGAGCCTGCGCGCTGCGGCGGCGGCGGCGGCGGCGGCGG 4080
Db 4040 CAGGACCGCTCAGAGCGGAGGAGTGTGAGCCTGCGCGCTGCGGCGGCGGCGGCGGCGGCGGCGG 4099
QY 4081 GCGAGCGGCGGAGTGTCTCATCAACAGAGTGTGAGCCTGCGCGCTGCGGCGGCGGCGGCGGCGG 4140
Db 4100 GCGAGCGGCGGAGTGTCTCATCAACAGAGTGTGAGCCTGCGCGCTGCGGCGGCGGCGGCGGCGG 4159
QY 4141 GGCTCCGAGCGGAGTGTCTCATCAACAGAGTGTGAGCCTGCGCGCTGCGGCGGCGGCGGCGGCGG 4200
Db 4160 GGCTCCGAGCGGAGTGTCTCATCAACAGAGTGTGAGCCTGCGCGCTGCGGCGGCGGCGGCGGCGG 4219
QY 4201 AGCAGTCCATCGGCGGCGGAGTGTGAGCCTGCGCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGG 4260
Db 4220 AGCAGTCCATCGGCGGCGGAGTGTGAGCCTGCGCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGG 4279
QY 4261 TATTTTGTGTCAGCGGCGGAGTGTGAGCCTGCGCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGG 4320
Db 4280 TATTTTGTGTCAGCGGCGGAGTGTGAGCCTGCGCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGG 4339
QY 4321 CACGAGTATGTACGCGGAGCCCGCAGCGTGCCTCTCAATTTTCATAGCCCGCGGCGGCTTCC 4380

Db 4340 CACGAGTATGTACGCGGAGCCCGCAGCGTGCCTCTCAATTTTCATAGCCCGCGGCGGCTTCC 4399
QY 4381 CAGCATGGCCCTTTTCAAGGATCGCATGGCATGGCATGGCATGGCATGGCATGGCATGGCATGG 4440
Db 4400 CAGCATGGCCCTTTTCAAGGATCGCATGGCATGGCATGGCATGGCATGGCATGGCATGGCATGG 4459
QY 4441 ATGGGGGGCGGGGGCGGGGGTCCCTGTATACGACCGGAAACCAAGTCTTCAAAATTTCCGGCC 4500
Db 4460 ATGGGGGGCGGGGGCGGGGGTCCCTGTATACGACCGGAAACCAAGTCTTCAAAATTTCCGGCC 4519
QY 4501 AGCAGTCTGTTCAGGACGAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 4560
Db 4520 AGCAGTCTGTTCAGGACGAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 4579
QY 4561 CCGGCCAGGACCGCTCCCTGTATACGACCGGAAACCAAGTCTTCAAAATTTCCGGCC 4620
Db 4580 CCGGCCAGGACCGCTCCCTGTATACGACCGGAAACCAAGTCTTCAAAATTTCCGGCC 4639
QY 4621 ACTCGGAGACCGTACAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 4680
Db 4640 ACTCGGAGACCGTACAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 4699
QY 4681 AGCACCGAGTGTGTGACGAGCACTACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 4740
Db 4700 AGCACCGAGTGTGTGACGAGCACTACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 4759
QY 4741 CTGATTTGAACTCGGAGTCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 4800
Db 4760 CTGATTTGAACTCGGAGTCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 4819
QY 4801 CTGTGCGGAGGAGCAGTGCCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 4860
Db 4820 CTGTGCGGAGGAGCAGTGCCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 4879
QY 4861 TTCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 4920
Db 4880 TTCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 4939
QY 4921 TTCTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 4980
Db 4940 TTCTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 4999
QY 4981 TTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 5040
Db 5000 TTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 5059
QY 5041 GCGAGGCGTGGGAGAACTTTGTA 5063
Db 5060 GCGAGGCGTGGGAGAACTTTGTA 5082

RESULT 5

US-09-949-016-2457
; Sequence 2457, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2457
; LENGTH: 5103
; TYPE: DNA

QY 2104 TCCAAACACACATCTACTGACAGACGTCAGCCTGAAGACCATCAGCCGCGCTTCATG 2163
Db 2101 TCCAAACACACATCTACTGACAGACGTCAGCCTGAAGACCATCAGCCGCGCTTCATG 2160
QY 2164 AACGGAGCTCGGTGGAGCAAGTGTGAGTTTGGCTTGTACTACCCGAGGCGATGGCC 2223
Db 2161 AACGGAGCTCGGTGGAGCAAGTGTGAGTTTGGCTTGTACTACCCGAGGCGATGGCC 2220
QY 2224 GTTGACTGGATGGGCAAGAACTCTACTGGGCGGACACTGGGACCAACAGATCGAAGTG 2283
Db 2221 GTTGACTGGATGGGCAAGAACTCTACTGGGCGGACACTGGGACCAACAGATCGAAGTG 2280
QY 2284 CGCGGCTGGAGCGGCGAGTTCGGCAAGTCTCTGTGTGGAGGAGCTTGGAACAACCCGAGG 2343
Db 2281 CGCGGCTGGAGCGGCGAGTTCGGCAAGTCTCTGTGTGGAGGAGCTTGGAACAACCCGAGG 2340
QY 2344 TCGCTGGCCCTGGATCCCAACGAGGCTACATCTACTGGACCGAGTGGGGCGGCAAGCCG 2403
Db 2341 TCGCTGGCCCTGGATCCCAACGAGGCTACATCTACTGGACCGAGTGGGGCGGCAAGCCG 2400
QY 2404 AGGATCGTGGCGGCTTCATGACGGGACCAACTGCATGACGCTGGTGGAACAAGTGGGC 2463
Db 2401 AGGATCGTGGCGGCTTCATGACGGGACCAACTGCATGACGCTGGTGGAACAAGTGGGC 2460
QY 2464 CGGGCCAAACGACCTCACCAATTGACTACGCTGACAGCGCCTCTACTGGACCGACCTGGAC 2523
Db 2461 CGGGCCAAACGACCTCACCAATTGACTACGCTGACAGCGCCTCTACTGGACCGACCTGGAC 2520
QY 2524 ACCAATGATGATGAGTGTCTCAACATGTCTGGGTGAGGAGCGGGTGTGATGTCGCAAGAT 2583
Db 2521 ACCAATGATGATGAGTGTCTCAACATGTCTGGGTGAGGAGCGGGTGTGATGTCGCAAGAT 2580
QY 2584 CTCGCGACCCGTTGCGTCTGACCCAGTACAGCGATATATATCTACTGACAGACTGGAT 2643
Db 2581 CTCGCGACCCGTTGCGTCTGACCCAGTACAGCGATATATATCTACTGACAGACTGGAT 2640
QY 2644 CTGCACAGCATGTAGCGGGCGGCAAGACTAGCGGCGGGAACCGCACCTCATCCAGGGC 2703
Db 2641 CTGCACAGCATGTAGCGGGCGGCAAGACTAGCGGCGGGAACCGCACCTCATCCAGGGC 2700
QY 2704 CACTGGACTTGTGATGGAATCTGTGTGTTCATCTCTCCCGCAGGATGGCTCAAT 2763
Db 2701 CACTGGACTTGTGATGGAATCTGTGTGTTCATCTCTCCCGCAGGATGGCTCAAT 2760
QY 2764 GACTGTATGCAACAACGCGGAGTGTGGGAGCTGTGCTTGCATCCCGCGGCGCAC 2823
Db 2761 GACTGTATGCAACAACGCGGAGTGTGGGAGCTGTGCTTGCATCCCGCGGCGCAC 2820
QY 2824 CGCTCGGCTCGGCTCACACTACACCTTGACCCCGAGCGGCAACTGCGAGCCGCGCC 2883
Db 2821 CGCTCGGCTCGGCTCACACTACACCTTGACCCCGAGCGGCAACTGCGAGCCGCGCC 2880
QY 2884 ACCACTTGTGTTGTCAGCCAGAAATCTGCCATCAGTCCGATGATCCCGGACGACGAG 2943
Db 2881 ACCACTTGTGTTGTCAGCCAGAAATCTGCCATCAGTCCGATGATCCCGGACGACGAG 2940
QY 2944 CACAGCCGGATCTCATCTGCGCTGCACTGAGCTGAGGACGTCGAAGCCATCGACTAT 3003
Db 2941 CACAGCCGGATCTCATCTGCGCTGCACTGAGCTGAGGACGTCGAAGCCATCGACTAT 3000
QY 3004 GACCCACTGGACAAGATTCATCTACTGGGTGGATGGGCGGCAAGAACATCAAGCGAGCCAAAG 3063
Db 3001 GACCCACTGGACAAGATTCATCTACTGGGTGGATGGGCGGCAAGAACATCAAGCGAGCCAAAG 3060
QY 3064 GACGACGGGACCCAGGAGCGGCTTGTGTTGACCTCTCTGAGGCCAAAGGCCAAACCCAGA 3119
Db 3061 GACGACGGGACCCAGGAGCGGCTTGTGTTGACCTCTCTGAGGCCAAAGGCCAAACCCAGA 3120
QY 3120 CAGGACGGGACCCAGGAGCGGCTTGTGTTGACCTCTCTGAGGCCAAAGGCCAAACCCAGA 3179
Db 3121 CAGGACGGGACCCAGGAGCGGCTTGTGTTGACCTCTCTGAGGCCAAAGGCCAAACCCAGA 3180

QY 3180 GGCCACCAATACCATCAACGCTCCACAGGCTGAGGGGGAAGCCATGGGGGTGGTCTGCG 3239
Db 3181 GGCCACCAATACCATCAACGCTCCACAGGCTGAGGGGGAAGCCATGGGGGTGGTCTGCG 3240
QY 3240 TGGGGAACGCGCAAGGCCAGGCGCATGCTGCTCAACGCGGAGGAGGGTACTCTGTACTT 3299
Db 3241 TGGGGAACGCGCAAGGCCAGGCGCATGCTGCTCAACGCGGAGGAGGGTACTCTGTACTT 3300
QY 3300 CACCAACATGACGAGACCGGCGAGCCAGATCGAAGCGGCGGCGGCGGCGGCGGCGG 3359
Db 3301 CACCAACATGACGAGACCGGCGAGCCAGATCGAAGCGGCGGCGGCGGCGGCGGCGG 3360
QY 3360 CGAGGCTCTCTTCAACACCGGCGCTCATCCGCGCTGTGGCCCTGTGGTGGCAACAACACT 3419
Db 3361 CGAGGCTCTCTTCAACACCGGCGCTCATCCGCGCTGTGGCCCTGTGGTGGCAACAACACT 3420
QY 3420 GGGCAAGCTGTTCTGGGTGGACCGGCGACTTGAAGCGCATTTGAGAGCTGTGACTGTACG 3479
Db 3421 GGGCAAGCTGTTCTGGGTGGACCGGCGACTTGAAGCGCATTTGAGAGCTGTGACTGTACG 3480
QY 3480 GGCCAAACGCGCTGACCTTGGAGGACGCCAATCGTGAGGCTCTGGGCGCTGACCATCT 3539
Db 3481 GGCCAAACGCGCTGACCTTGGAGGACGCCAATCGTGAGGCTCTGGGCGCTGACCATCT 3540
QY 3540 TGGCAAGCATCTTACTGGATCGACCGCGAGCAGATGATCGAGCGTGTGGAGAAAGAC 3599
Db 3541 TGGCAAGCATCTTACTGGATCGACCGCGAGCAGATGATCGAGCGTGTGGAGAAAGAC 3600
QY 3600 CACCGGGGAACGCGGACTCTGCATCCAGGGCGGTGTGCCCACTCTACTGGGATCCATGCG 3659
Db 3601 CACCGGGGAACGCGGACTCTGCATCCAGGGCGGTGTGCCCACTCTACTGGGATCCATGCG 3660
QY 3660 AGTCGAGGAGTCAAGCTTGGAGGAGTCTCAGCCACCCATGTCGCCGTGCAATGGTGG 3719
Db 3661 AGTCGAGGAGTCAAGCTTGGAGGAGTCTCAGCCACCCATGTCGCCGTGCAATGGTGG 3720
QY 3720 CTGCTCCACATCTGTATTGCAAGGGTGTGGGACACACCGGTGCTCATGCCAGTCCA 3779
Db 3721 CTGCTCCACATCTGTATTGCAAGGGTGTGGGACACACCGGTGCTCATGCCAGTCCA 3780
QY 3780 CCTCGTGTCTTGCAAACTCTGTGACTGTGGAGAGCGGCCACCTGCTCTCCGCGACA 3839
Db 3781 CCTCGTGTCTTGCAAACTCTGTGACTGTGGAGAGCGGCCACCTGCTCTCCGCGACA 3840
QY 3840 GTTTGCAATGTCACAGGGGAGTCAAGTGTATCCCGGGGCGCTGGCGCTGTGACGGCTT 3899
Db 3841 GTTTGCAATGTCACAGGGGAGTCAAGTGTATCCCGGGGCGCTGGCGCTGTGACGGCTT 3900
QY 3900 TCCGAGTGGCGATGACAGAGCGAGGAGGCTGCGCCGCTGTGCTCCGCGCGCCAGTT 3959
Db 3901 TCCGAGTGGCGATGACAGAGCGAGGAGGCTGCGCCGCTGTGCTCCGCGCGCCAGTT 3960
QY 3960 CCCCTGCGCGGGGTCAAGTGTGGACCTGCGCTGCGCTGCGAGCGGCGAGGAGACTG 4019
Db 3961 CCCCTGCGCGGGGTCAAGTGTGGACCTGCGCTGCGCTGCGAGCGGCGAGGAGACTG 4020
QY 4020 TCAGGACCGCTCAGACGAGCGGAGTGTGAGCCCATCTGCGCTGCCCAACACTGTCGGTG 4079
Db 4021 TCAGGACCGCTCAGACGAGCGGAGTGTGAGCCCATCTGCGCTGCCCAACACTGTCGGTG 4080
QY 4080 TGCAGCGGCGGAGTGTGCTCTCATCAACAGCAGTGGAGCTCTTCCCGGAGCTGTATCGA 4139
Db 4081 TGCAGCGGCGGAGTGTGCTCTCATCAACAGCAGTGGAGCTCTTCCCGGAGCTGTATCGA 4140
QY 4140 CGGCTCCGAGCAGTCAATGTGTGAAATCAACAGCGGCGCTCTCAGACGAGCGCGGCCA 4199
Db 4141 CGGCTCCGAGCAGTCAATGTGTGAAATCAACAGCGGCGCTCTCAGACGAGCGCGGCCA 4200
QY 4200 CAGCAGTGGCATCGGGCGGCTCATTTGGCATCATCTCTCTCTCTGTCATGGGTGGT 4259
Db 4201 CAGCAGTGGCATCGGGCGGCTCATTTGGCATCATCTCTCTCTCTCTGTCATGGGTGGT 4260
QY 4260 CTATTTTGTGTGCCAGCGGCTGTGTGTCAGCGCTATGCGGGGCGCAACGCGGCGCTTCCC 4319

Db 4261 CTAATTTGTGTCAGCGGTGTGTGTCAGCGCTATGCGGGGGCAACGGGCCCTTCCC 4320
QY 4320 GCACGAGTATGTACGGGACCCGACATGCGCCCTCAATTTTCATAGCCCCGGGGGTTTC 4379
Db 4321 GCACGAGTATGTACGGGACCCGACATGCGCCCTCAATTTTCATAGCCCCGGGGGTTTC 4380
QY 4380 CCAGCATGCCCTTTCACAGGATGTCATGCGGAAAGTTCATGATGAGCTTCGTCGAGCT 4439
Db 4381 CCAGCATGCCCTTTCACAGGATGTCATGCGGAAAGTTCATGATGAGCTTCGTCGAGCT 4440
QY 4440 GATGGGGGGCGGGGGGGTCCCTGTGACACCGGAAACACAGTCACAGGGGGCTTCGTC 4499
Db 4441 GATGGGGGGCGGGGGGGTCCCTGTGACACCGGAAACACAGTCACAGGGGGCTTCGTC 4500
QY 4500 CAGCAGCTGTCGACGAGGACGAGGACGAGGACGAGGACGAGGACGAGGACGAGGACG 4559
Db 4501 CAGCAGCTGTCGACGAGGACGAGGACGAGGACGAGGACGAGGACGAGGACGAGGACG 4560
QY 4560 CCGGCGCAGGACCCCTCCCTGTGACAAATGACAGATGTTCTACTTTCAAACATTCGGC 4619
Db 4561 CCGGCGCAGGACCCCTCCCTGTGACAAATGACAGATGTTCTACTTTCAAACATTCGGC 4620
QY 4620 CACTCGGAGACCGTACAGGACCCCTACATCATTCGAGGAATGGCGCCCGGACGAGGCCCTG 4679
Db 4621 CACTCGGAGACCGTACAGGACCCCTACATCATTCGAGGAATGGCGCCCGGACGAGGCCCTG 4680
QY 4680 CAGCACCGAGCTGTGTGACAGGACGACGACGACGACGACGACGACGACGACGACGACG 4739
Db 4681 CAGCACCGAGCTGTGTGACAGGACGACGACGACGACGACGACGACGACGACGACGACG 4740
QY 4740 CTTGATTTGAACTCGGACTCAGACCCCTATCCACCCCTATCCACCCCTATCCACCCCTATCC 4799
Db 4741 CTTGATTTGAACTCGGACTCAGACCCCTATCCACCCCTATCCACCCCTATCCACCCCTATCC 4800
QY 4800 CTTGTCGCGGAGGACGACGTCGCGCCCTCGCGCCCTCGCGCCCTCGCGCCCTCGCGCCCT 4859
Db 4801 CTTGTCGCGGAGGACGACGTCGCGCCCTCGCGCCCTCGCGCCCTCGCGCCCTCGCGCCCT 4860
QY 4860 CTTCCCGCCCTCGCTCCCTCGCTGACGAGCTATCTGACGCTGCGCGCGGCGCTCTGCG 4919
Db 4861 CTTCCCGCCCTCGCTCCCTCGCTGACGAGCTATCTGACGCTGCGCGCGGCGCTCTGCG 4920
QY 4920 CTTCTCTGCGCCCTGTAATAGTTTAAATATGAACAAAGAAATATATTTATGA 4979
Db 4921 CTTCTCTGCGCCCTGTAATAGTTTAAATATGAACAAAGAAATATATTTATGA 4980
QY 4980 TTTAAATAAATAAATAATTTGGATTTTAAATAACATGAGAAATGTGAATGATGGGGT 5039
Db 4981 TTTAAATAAATAAATAATTTGGATTTTAAATAACATGAGAAATGTGAATGATGGGGT 5040
QY 5040 GGGCAGGGCTGGGAGAACTTTGTA 5063
Db 5041 GGGCAGGGCTGGGAGAACTTTGTA 5064

RESULT 6

US-09-060-299-1
; Sequence 1, Application US/09060299
; Patent No. 6545137
; GENERAL INFORMATION:
; APPLICANT: Todd, John A
; APPLICANT: Hess, John W
; APPLICANT: Caskey, Charles T
; APPLICANT: Cox, Roger D
; APPLICANT: Gerhold, David
; APPLICANT: Hammond, Holly
; APPLICANT: Hey, Patricia
; APPLICANT: Kawaguchi, Yoshihiko
; APPLICANT: Merriman, Tony R
; APPLICANT: Metzker, Michael L
; TITLE OF INVENTION: No. 6545137el Receptor
; NUMBER OF SEQUENCES: 455

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon and Vanderhye
; STREET: 1100 No. 6545137th Glebe Road, Eighth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: US
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA: US/09/060,299
; APPLICATION NUMBER: US/09/060,299
; FILING DATE: 15-APR-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/043,553
; FILING DATE: 15-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/048,740
; FILING DATE: 05-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: B. J. Sadoff
; REGISTRATION NUMBER: 36,663
; REFERENCE/DOCKET NUMBER: 620-35
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)816-4091
; TELEFAX: (703)816-4100
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5098 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-060-299-1

Query Match 99.3%; Score 5028; DB 4; Length 5098;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 5053; Conservative 0; Mismatches 5; Indels 2; Gaps 2;

QY 4 ATGAGAGCCCGAGTGAAGCG 63
Db 1 ATGAGAGCCCGAGTGAAGCG 60
QY 64 CCG 123
Db 61 CCG 120
QY 124 TGCCCG 183
Db 121 TGCCCG 180
QY 184 CTGCTGGACCG 243
Db 181 CTGCTGGACCG 240
QY 244 GCG 303
Db 241 GCG 300
QY 304 GAGGCCATCAAGCAGACCTTACCTGAACACGACGCGGGGCGCGCGCGCGCGCGCGCGCGCG 363
Db 301 GAGGCCATCAAGCAGACCTTACCTGAACACGACGCGGGGCGCGCGCGCGCGCGCGCGCGCG 360
QY 364 TCG 423
Db 361 TCG 420
QY 424 ACGGACTCAGAGACCAACCGCATCGAGTGGCGCAACCTCAATGTCACATCCCGAAGGTG 483
Db 421 ACGGACTCAGAGACCAACCGCATCGAGTGGCGCAACCTCAATGTCACATCCCGAAGGTG 480
QY 484 CTCTTTCGCGAGGACCTTTGACCGCCTAGGGCCCATCGCCTTGGACCCCGCTCACGGGTAC 543

Db 481 |||||CTCTTCTGGCAGGACCTTTGACCCAGCCGAGGCGCATCGCCTTTGGACCCCGCTCA CGGGTAC 540
QY 544 ATGTACTGGACAGACTGGGTGGAGACGCCCGCGATTGAGCGGGCAGGATGGATGGCAGC 603
Db 541 ATGTACTGGACAGACTGGGTGGAGACGCCCGCGATTGAGCGGGCAGGATGGATGGCAGC 600
QY 604 ACCCGGAAGATCATTTGTGGACTCGGACATTTACTGGCCCAATGGACTGACCATCGACCTG 663
Db 601 ACCCGGAAGATCATTTGTGGACTCGGACATTTACTGGCCCAATGGACTGACCATCGACCTG 660
QY 664 GAGGAGCAGAGCTCTACTGGGTGACGCCAAGCTCAGCTTCATCCACCTGTCGCAACCTG 723
Db 661 GAGGAGCAGAGCTCTACTGGGTGACGCCAAGCTCAGCTTCATCCACCTGTCGCAACCTG 720
QY 724 GACGCTCGTTCCGGCAGAGGTGTGGAGGGCAGCCTGACGCCACCCCTTCGCCCTGACG 783
Db 721 GACGCTCGTTCCGGCAGAGGTGTGGAGGGCAGCCTGACGCCACCCCTTCGCCCTGACG 780
QY 784 CTCTCCGGGACACTCTGTACTGGACAGACTGGCAGACCCCGCTCCATCCATGCTGCAAC 843
Db 781 CTCTCCGGGACACTCTGTACTGGACAGACTGGCAGACCCCGCTCCATCCATGCTGCAAC 840
QY 844 AAGCGCACTGGGGGAGAGGAAAGAGATCTGTAGTGCCTCTACTCAACCCATGAGCATC 903
Db 841 AAGCGCACTGGGGGAGAGGAAAGAGATCTGTAGTGCCTCTACTCAACCCATGAGCATC 900
QY 904 CAGGTGCTGACCGAGGAGCGGCGCTTTCTTCCACACTGCTGTGAGGAGACATGGC 963
Db 901 CAGGTGCTGACCGAGGAGCGGCGCTTTCTTCCACACTGCTGTGAGGAGACATGGC 960
QY 964 GECTGTCTCCACCTGTGCTGTCTCCCAAGCAGAGCCTTTCTACACATGCGCTGCCCC 1023
Db 961 GECTGTCTCCACCTGTGCTGTCTCCCAAGCAGAGCCTTTCTACACATGCGCTGCCCC 1020
QY 1024 ACGGGTGTGAGCTGACGAGCAAAACGGCAGGACGTGTAAAGCAGGAGCGGAGAGTGCTG 1083
Db 1021 ACGGGTGTGAGCTGACGAGCAAAACGGCAGGACGTGTAAAGCAGGAGCGGAGAGTGCTG 1080
QY 1084 CTGCTGGCCCGGCGAGCGGACCTACGAGGATCTCGCTGGACACGCCGGACTTACCGAC 1143
Db 1081 CTGCTGGCCCGGCGAGCGGACCTACGAGGATCTCGCTGGACACGCCGGACTTACCGAC 1140
QY 1144 ATCGTGTGAGGTGGACGACATCCGGCACGCCATTGGCATCGACTAGACCCGCTAGAG 1203
Db 1141 ATCGTGTGAGGTGGACGACATCCGGCACGCCATTGGCATCGACTAGACCCGCTAGAG 1200
QY 1204 GGCTATGTCTACTGGACAGATGACGAGGTGCGGGCCATTCGCGAGGGCGTACTTGGACGGG 1263
Db 1201 GGCTATGTCTACTGGACAGATGACGAGGTGCGGGCCATTCGCGAGGGCGTACTTGGACGGG 1260
QY 1264 TCTGGGGCGCAGACCTGGTCAACACCGAGATCAACGACCCCGATGGCATCGCGGTGAC 1323
Db 1261 TCTGGGGCGCAGACCTGGTCAACACCGAGATCAACGACCCCGATGGCATCGCGGTGAC 1320
QY 1324 TGGGTGGCCCAAACTCTACTGGACGACACGGGCGACGCGACCGCATCGAGGTGACGCGC 1383
Db 1321 TGGGTGGCCCAAACTCTACTGGACGACACGGGCGACGCGACCGCATCGAGGTGACGCGC 1380
QY 1384 CTC AACCGCACCTCCCGCAAGATCCTGTGTGTCGAGGACCTTGGACGAGCCCGGAGCCATC 1443
Db 1381 CTC AACCGCACCTCCCGCAAGATCCTGTGTGTCGAGGACCTTGGACGAGCCCGGAGCCATC 1440
QY 1444 GCATGCAACCCCGTATGGGCTCATGTACTGGACAGACTGGGAGAGAACCTTAAATC 1503
Db 1441 GCATGCAACCCCGTATGGGCTCATGTACTGGACAGACTGGGAGAGAACCTTAAATC 1500
QY 1504 GAGTGTGCAACTTGGATGGGCGAGCGGCTGTGTGTCGTCATGCTCCCTCGGTTGG 1563
Db 1501 GAGTGTGCAACTTGGATGGGCGAGGAGCGGCTGTGTGTCGTCATGCTCCCTCGGTTGG 1560
QY 1564 CCCAACCGCTTGGCCTGGACCTTCGAGGAGGGAGCTCTACTGGGAGAGCCCAAGACA 1623

Db 1561 CCCAACCGCTTGGCCTTGGACCTTCGAGGAGGGAGCTCTACTGGGGAGAGCCCAAGACA 1620
QY 1624 GACAAGATCGAGGTGATCAATGTTGATGGGACGAGAGCGCGGACCTCTCTGGAGACAG 1683
Db 1621 GACAAGATCGAGGTGATCAATGTTGATGGGACGAGAGCGCGGACCTCTCTGGAGACAG 1680
QY 1684 CTC CGGCACATTTTTCGGGTTCA CGCTGTCTGGGGACTTCACTACTGACTGACTGGCAG 1743
Db 1681 CTC CGGCACATTTTTCGGGTTCA CGCTGTCTGGGGACTTCACTACTGACTGACTGGCAG 1740
QY 1744 CGCGCAGCATCGAGCGGGTGCAAAAGTCAAGCGCAGCCGCGGAGCTCATATTGACAG 1803
Db 1741 CGCGCAGCATCGAGCGGGTGCAAAAGTCAAGCGCAGCCGCGGAGCTCATATTGACAG 1800
QY 1804 CTGCGCGACCTGATGGGGCTCAAAAGCTGTGAAATGTGGCCAAAGGTGCTCGGACCAACCCG 1863
Db 1801 CTGCGCGACCTGATGGGGCTCAAAAGCTGTGAAATGTGGCCAAAGGTGCTCGGACCAACCCG 1860
QY 1864 TGTGCGGACAGGAAACGGGGGGTGAGCCACCTGTGCTCTTTCACACCCCGACCAACCCGG 1923
Db 1861 TGTGCGGACAGGAAACGGGGGGTGAGCCACCTGTGCTCTTTCACACCCCGACCAACCCGG 1920
QY 1924 TGTGCTGCCCCATCGGCTGGAGCTGTGTGATGACATGAAGACTGTGATCTGCTGAG 1983
Db 1921 TGTGCTGCCCCATCGGCTGGAGCTGTGTGATGACATGAAGACTGTGATCTGCTGAG 1980
QY 1984 GCCTTCTTGGTCTTTCACAGCAGAGCGGCCATCCACAGGATCTCCCTCGAGACCAATAAC 2043
Db 1981 GCCTTCTTGGTCTTTCACAGCAGAGCGGCCATCCACAGGATCTCCCTCGAGACCAATAAC 2040
QY 2044 AACGACGTGGCCATCCCGCTCACGGGGCTCAAGGAGGCTCAGGCCCTGGACTTTGATGTG 2103
Db 2041 AACGACGT-GCCATCCCGCTCACGGGGCTCAAGGAGGCTCAGGCCCTGGACTTTGATGTG 2099
QY 2104 TCCAAACCAACATCTACTGGAAGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAG 2163
Db 2100 TCCAAACCAACATCTACTGGAAGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAG 2159
QY 2164 AACGGGAGCTCGGTGGAGCAGCTGTGGAGTTTGGCCTTGTGACTACCCGAGGGCATGCC 2223
Db 2160 AACGGGAGCTCGGTGGAGCAGCTGTGGAGTTTGGCCTTGTGACTACCCGAGGGCATGCC 2219
QY 2224 GTTGACTTGGATGGGCAAGAACCTCTACTGGGCCGACACTGGGACCAACAGAAATCGAAGTG 2283
Db 2220 GTTGACTTGGATGGGCAAGAACCTCTACTGGGCCGACACTGGGACCAACAGAAATCGAAGTG 2279
QY 2284 GCGCGCTGGAAGCGGCTTCCGGCAAGTCTCTGTGTGGAGGACTTGGACAAACCCGAGG 2343
Db 2280 GCGCGCTGGAAGCGGCTTCCGGCAAGTCTCTGTGTGGAGGACTTGGACAAACCCGAGG 2339
QY 2344 TCGCTGGCCCTGGATCCCAACGAGGCTCATCTACTGGACCGGAGTGGGGCGGCAAGCCG 2403
Db 2340 TCGCTGGCCCTGGATCCCAACGAGGCTCATCTACTGGACCGGAGTGGGGCGGCAAGCCG 2399
QY 2404 AGGATCTGTGGGGCTTTCATGGACGGGACCAACCTGTGATGAGCTGGTGGACAAAGTGGGC 2463
Db 2400 AGGATCTGTGGGGCTTTCATGGACGGGACCAACCTGTGATGAGCTGGTGGACAAAGTGGGC 2459
QY 2464 CGGGCCAAACGACTTCAACATTTGACTAGCGTGTGAGCGGCTTCTACTGGACCGACTGGAC 2523
Db 2460 CGGGCCAAACGACTTCAACATTTGACTAGCGTGTGAGCGGCTTCTACTGGACCGACTGGAC 2519
QY 2524 ACCAATGATGATCGATCGTCCACATCTGGGTGAGGAGGGGTGCTGATTTGCCGACAT 2583
Db 2520 ACCAATGATGATCGATCGTCCACATCTGGGTGAGGAGGGGTGCTGATTTGCCGACAT 2579
QY 2584 CTC CGGCACCCGTTTCGGTCTGACCGAGTACAGCAATATATCTACTTGGACAGACTGGAAT 2643
Db 2580 CTC CGGCACCCGTTTCGGTCTGACCGAGTACAGCAATATATCTACTTGGACAGACTGGAAT 2639
QY 2644 CTGCAAGCATTTGAGCGGGCGGACAGACTAGCGGCGGAAACCGCAACCTCATCCAGGGC 2703
Db 2640 CTGCAAGCATTTGAGCGGGCGGACAGACTAGCGGCGGAAACCGCAACCTCATCCAGGGC 2699

QY	2704	CACCTGGACTTCGTGATGAGACATCTCTGGTGTTCACACTCTCTCCGCGCAGGATGGCCTCAAT	2763
Db	2700	CACCTGGACTTCGTGATGAGACATCTCTGGTGTTCACACTCTCTCCGCGCAGGATGGCCTCAAT	2759
QY	2764	GACTGTATGCAAAACAAACGGGCGAGTGTGGGCACTGTGCTTGCATCCCGCGGCGCCAC	2823
Db	2760	GACTGTATGCAAAACAAACGGGCGAGTGTGGGCACTGTGCTTGCATCCCGCGGCGCCAC	2819
QY	2824	CGCTGGGTGGCCTTCACTACACCTTGACCCAGCAGCCGCAATGTCAGCCCGCC	2883
Db	2820	CGCTGGGTGGCCTTCACTACACCTTGACCCAGCAGCCGCAATGTCAGCCCGCC	2879
QY	2884	ACCACTTCTTCTGCTTTCAGCCAGAAATCTGCCATCAGTCGGATGATCCCGAGACACCA	2943
Db	2880	ACCACTTCTTCTGCTTTCAGCCAGAAATCTGCCATCAGTCGGATGATCCCGAGACACCA	2939
QY	2944	CACAGCCGGATCTCATCTCTGCCCTGATGAGTCTGAGGAACTGTCAGGACCTCACTAT	3003
Db	2940	CACAGCCGGATCTCATCTCTGCCCTGATGAGTCTGAGGAACTGTCAGGACCTCACTAT	2999
QY	3004	GACCACTGGACAACTTCACTACCTGGGTGAGTGGCGCGAGAAATCAAGCGACCAAG	3063
Db	3000	GACCACTGGACAACTTCACTACCTGGGTGAGTGGCGCGAGAAATCAAGCGACCAAG	3059
QY	3064	GACGACGGACCCAGCCCTTTGTTTGTGACCTCTCTGAGCCAGGCGAAACCCAGACAGG	3123
Db	3060	GACGACGGACCCAGCCCTTTGTTTGTGACCTCTCTGAGCCAGGCGAAACCCAGACAGG	3119
QY	3124	CAGCCCCACGACCTCAGATCGACATCTACAGCCGAGCACTGTCTCGACGTGCGAGGCC	3183
Db	3120	CAGCCCCACGACCTCAGATCGACATCTACAGCCGAGCACTGTCTCGACGTGCGAGGCC	3179
QY	3184	ACCAATACATCAAGCTTCAAGGTGAGCGGGGAGCCATGCGGGTGGTCTGGTGGG	3243
Db	3180	ACCAATACATCAAGCTTCAAGGTGAGCGGGGAGCCATGCGGGTGGTCTGGTGGG	3239
QY	3244	GACCGCGAAGACCCAGCGCCATCTCTGTCAGCGGAGCGAGGTACTCTGACTTCAAC	3303
Db	3240	GACCGCGAAGACCCAGCGCCATCTCTGTCAGCGGAGCGAGGTACTCTGACTTCAAC	3299
QY	3304	AACATGCAAGACCGGGCAGCAGAACTGAAACCGCAGCCCTGGAAGCAGCGAGCGCAG	3363
Db	3300	AACATGCAAGACCGGGCAGCAGAACTGAAACCGCAGCCCTGGAAGCAGCGAGCGCAG	3359
QY	3364	GTCTCTTTCACACCGGCTCATCCGCTCTGCGCTTGGCTTGGTGGACAAACACTGGGC	3423
Db	3360	GTCTCTTTCACACCGGCTCATCCGCTCTGCGCTTGGCTTGGTGGACAAACACTGGGC	3419
QY	3424	AAGCTGTTCTGGGTGAGCGCGACCTGAAGCGCATTTGAGAGCTGTGACCTGTGAGGGGCC	3483
Db	3420	AAGCTGTTCTGGGTGAGCGCGACCTGAAGCGCATTTGAGAGCTGTGACCTGTGAGGGGCC	3479
QY	3484	AACCGCTGACCTGGAGAGCGCAACATCTGACGCTCTGGGCTTGACCATCTTTGGC	3543
Db	3480	AACCGCTGACCTGGAGAGCGCAACATCTGACGCTCTGGGCTTGACCATCTTTGGC	3539
QY	3544	AAGCATCTTACTGATCGACCGCCAGCAGATGATCGAGCTGTGAGAGAACCAACC	3603
Db	3540	AAGCATCTTACTGATCGACCGCCAGCAGATGATCGAGCTGTGAGAGAACCAACC	3599
QY	3604	GGGACAAAGCGACTCGCATCCAGGGCGTGTGCGCCACCTCACTGGGATCCATCGAGTG	3663
Db	3600	GGGACAAAGCGACTCGCATCCAGGGCGTGTGCGCCACCTCACTGGGATCCATCGAGTG	3659
QY	3664	GAGGAAGTCAGCTGGAGAGTTCTCAGCCCAACCATGTGCGGCTGACAAATGTTGGTGC	3723
Db	3660	GAGGAAGTCAGCTGGAGAGTTCTCAGCCCAACCATGTGCGGCTGACAAATGTTGGTGC	3719
QY	3724	TCCACATCTGATTTGCCAAGGTGATGGGACACACAGGTGCTCATGCCAGTCCACTC	3783
Db	3720	TCCACATCTGATTTGCCAAGGTGATGGGACACACAGGTGCTCATGCCAGTCCACTC	3779

QY	3784	GTGCTCTTGCAAACTGCTGACCTGTGGAGAGCGCGCCACCTGCTCCCGGACAGTTT	3843
Db	3780	GTGCTCTTGCAAACTGCTGACCTGTGGAGAGCGCGCCACCTGCTCCCGGACAGTTT	3839
QY	3844	GCATGTGCAACAGGGAGATCGACTGTATCCCGGGGCTGGCGCTGTGAGCGCTTCCC	3903
Db	3840	GCATGTGCAACAGGGAGATCGACTGTATCCCGGGGCTGGCGCTGTGAGCGCTTCCC	3899
QY	3904	GAGTGCATGACACAGAGCGACGAGAGGGCTGCGCCCGTGTGCTCCGCGCCAGTTCGCC	3963
Db	3900	GAGTGCATGACACAGAGCGACGAGAGGGCTGCGCCCGTGTGCTCCGCGCCAGTTCGCC	3959
QY	3964	TGCGCGGGGTGAGTGTGTGGACCTGCGCTGCGCTGCGAGCGGAGGAGACTGTGAG	4023
Db	3960	TGCGCGGGGTGAGTGTGTGGACCTGCGCTGCGCTGCGAGCGGAGGAGACTGTGAG	4019
QY	4024	GACCGCTCAGACGAGGGGAGTGTGAACCTCTGCTGCGCCAAACAGTTCGGGTGTGCG	4083
Db	4020	GACCGCTCAGACGAGGGGAGTGTGAACCTCTGCTGCGCCAAACAGTTCGGGTGTGCG	4079
QY	4084	AGCGGCGAGTGTGCTCATCAAAACAGAGTGGAGTCTCTTCCCGGACTGTATCGACGCG	4143
Db	4080	AGCGGCGAGTGTGCTCATCAAAACAGAGTGGAGTCTCTTCCCGGACTGTATCGACGCG	4138
QY	4144	TCCGACGAGCTCATGTGTGAATCACCAGCGCCCTCAGACGACAGCCCGGCCACAGC	4203
Db	4139	TCCGACGAGCTCATGTGTGAATCACCAGCGCCCTCAGACGACAGCCCGGCCACAGC	4198
QY	4204	AGTGCCATCGGCGCCGCTCATTTGGCATCATCTCTCTCTCTTTCGTGCTGCTGCTAT	4263
Db	4199	AGTGCCATCGGCGCCGCTCATTTGGCATCATCTCTCTCTCTTTCGTGCTGCTGCTAT	4258
QY	4264	TTTGTGTGCGAGCGCTGTGTGCGAGGCTATGCGGGGCGCAACGGGCGCTTCCCGGAC	4323
Db	4259	TTTGTGTGCGAGCGCTGTGTGCGAGGCTATGCGGGGCGCAACGGGCGCTTCCCGGAC	4318
QY	4324	GAGTATGTGAGCGGACCCCGCAGTCCCTCAATTTCTAGAGCCCGGGGCTTCCCGAG	4383
Db	4319	GAGTATGTGAGCGGACCCCGCAGTCCCTCAATTTCTAGAGCCCGGGGCTTCCCGAG	4378
QY	4384	CATGGCCCTTTCACAGGATCGCATGCGGAAAGTCCATGATGAGCTCCGTGAGCTGATG	4443
Db	4379	CATGGCCCTTTCACAGGATCGCATGCGGAAAGTCCATGATGAGCTCCGTGAGCTGATG	4438
QY	4444	GGGGGCGGGGCGGGGCTGCTGTAACGAGAACCACTGACAGGGGCTTCTTCAGC	4503
Db	4439	GGGGGCGGGGCGGGGCTGCTGTAACGAGAACCACTGACAGGGGCTTCTTCAGC	4498
QY	4504	AGCTGCTCCAGCACAAAGGCGACGCTGTACCGCGCGATCTGAAACCGCGCCCTCCCGG	4563
Db	4499	AGCTGCTCCAGCACAAAGGCGACGCTGTACCGCGCGATCTGAAACCGCGCCCTCCCGG	4558
QY	4564	GCCACGACCCCTTCTGTACAACATGGAATGTTTACTTCTTCAAAACATTCGCGGCACT	4623
Db	4559	GCCACGACCCCTTCTGTACAACATGGAATGTTTACTTCTTCAAAACATTCGCGGCACT	4618
QY	4624	GCGAGACGCTACAGGCGCTCATCATTTGAGGAAATGCGGCGCCCGCGAGCGCCCTGAGC	4683
Db	4619	GCGAGACGCTACAGGCGCTCATCATTTGAGGAAATGCGGCGCCCGCGAGCGCCCTGAGC	4678
QY	4684	ACCGACGCTGTGACACGAGCTACAGCGCGAGCGCTGGAAGGCGAGCAAGTACTACTG	4743
Db	4679	ACCGACGCTGTGACACGAGCTACAGCGCGAGCGCTGGAAGGCGAGCAAGTACTACTG	4738
QY	4744	GATTTGAATCGGACTCAGACCCCTATFCAACCCCAACCAAGCGGCGGAGCAAGTACTG	4803
Db	4739	GATTTGAATCGGACTCAGACCCCTATFCAACCCCAACCAAGCGGCGGAGCAAGTACTG	4798
QY	4804	TCCGCGAGGACAGCTGCGCGCCCTGCGCGCCACCGAGAGGAGTACTTCCATCTCTTC	4863
Db	4799	TCCGCGAGGACAGCTGCGCGCCCTGCGCGCCACCGAGAGGAGTACTTCCATCTCTTC	4858
QY	4864	CCGCCCCCTCCGCTCCGCTGACGAGACTCATCTGAGCTCATCTGCGCGGGGCGCACTCTGGGTTTC	4923

QY 1084 CTGCTGGCCCGCGGACGGACCTACGAGGATCTCGCTGGACACGCGCGGACTTACCGAC 1143
DB 1081 CTGCTGGCCCGCGGACGGACCTACGAGGATCTCGCTGGACACGCGCGGACTTACCGAC 1140
QY 1144 ATCGTGCTGCAAGTGGACGACATCCGGCAGCGCATTTGCCATCGACTAGCAGCCGCTAGAG 1203
DB 1141 ATCGTGCTGCAAGTGGACGACATCCGGCAGCGCATTTGCCATCGACTAGCAGCCGCTAGAG 1200
QY 1204 GCTATGTCTACTGACACAGATGACGAGTGGGGCCATCCCGCAGCGGCTACTGAGACCGG 1263
DB 1201 GCTATGTCTACTGACACAGATGACGAGTGGGGCCATCCCGCAGCGGCTACTGAGACCGG 1260
QY 1264 TCTGGGGCGCAGACCTGGTCAACACCGAGATCAACGACCGGATCGGATGGGATCGCGGTGCAC 1323
DB 1261 TCTGGGGCGCAGACCTGGTCAACACCGAGATCAACGACCGGATCGGATGGGATCGCGGTGCAC 1320
QY 1324 TGGGTGGCCCGAAACCTTCTACTGACCGACACGCGCAGCGGATCGAGGTGACGCGC 1383
DB 1321 TGGGTGGCCCGAAACCTTCTACTGACCGACACGCGCAGCGGATCGAGGTGACGCGC 1380
QY 1384 CTCACGGGACCTCCGCAAGATCTGTGTGCGAGGACCTGAGACGCGCCCGAGCCATC 1443
DB 1381 CTCACGGGACCTCCGCAAGATCTGTGTGCGAGGACCTGAGACGCGCCCGAGCCATC 1440
QY 1444 GCACGTGACCCCGTGCATGGGCTCATGTACTGACAGACTGGGAGAGAACCTTAATATC 1503
DB 1441 GCACGTGACCCCGTGCATGGGCTCATGTACTGACAGACTGGGAGAGAACCTTAATATC 1500
QY 1504 GAGTGTGCAACTTTGGATGGGCGAGCGGGGTGTGCTGGTCAATGCCCTCCCTCGGGTGG 1563
DB 1501 GAGTGTGCAACTTTGGATGGGCGAGCGGGGTGTGCTGGTCAATGCCCTCCCTCGGGTGG 1560
QY 1564 CCCAAGCGCTGGCCCTGGACCTGACGAGGAGGAGGCTCTACTGAGGAGACGCCAAGACA 1623
DB 1561 CCCAAGCGCTGGCCCTGGACCTGACGAGGAGGAGGCTCTACTGAGGAGACGCCAAGACA 1620
QY 1624 GACAAGATCGAGGTGATCAATGTTGATGCGGACGAGGCGGACCCCTCTGAGGAGACAAG 1683
DB 1621 GACAAGATCGAGGTGATCAATGTTGATGCGGACGAGGCGGACCCCTCTGAGGAGACAAG 1680
QY 1684 CTCGCGCACATTTTCGGGTTCACTGCTGGGGGACTTCACTACTGAGACTGACTGGCAG 1743
DB 1681 CTCGCGCACATTTTCGGGTTCACTGCTGGGGGACTTCACTACTGAGACTGACTGGCAG 1740
QY 1744 CGCGCAGCATCGACGGGTGACAGGTCAAGCCAGCGGAGCTCATATTGACCAAG 1803
DB 1741 CGCGCAGCATCGACGGGTGACAGGTCAAGCCAGCGGAGCTCATATTGACCAAG 1800
QY 1804 CTGCGCGACCTGATGGGGCTCAAGCTGTGAATGTGGCCAAAGTCTGCGGAACCAACCCG 1863
DB 1801 CTGCGCGACCTGATGGGGCTCAAGCTGTGAATGTGGCCAAAGTCTGCGGAACCAACCCG 1860
QY 1864 TGTGGGACAGGAACGGGGGTGAGCCACCTGTGCTTCTTCAACCCCGACGCAACCCGG 1923
DB 1861 TGTGGGACAGGAACGGGGGTGAGCCACCTGTGCTTCTTCAACCCCGACGCAACCCGG 1920
QY 1924 TGTGGTCCCCCATCGGCTGGAGCTGTGATGACATGAAGACCTGATCTGCTGCTGAG 1983
DB 1921 TGTGGTCCCCCATCGGCTGGAGCTGTGATGACATGAAGACCTGATCTGCTGCTGAG 1980
QY 1984 GCCTTCTTGGTCTTCAACGACAGAGCGCCCATCAACAGGATCTCCCTGAGACCAATAAC 2043
DB 1981 GCCTTCTTGGTCTTCAACGACAGAGCGCCCATCAACAGGATCTCCCTGAGACCAATAAC 2040
QY 2044 AACGACGTGGCCATCCCGCTCACGGGGGTCAAGGAGGCTCAGGCCCTGGACTTTGATGTG 2103
DB 2041 AACGACGTGGCCATCCCGCTCACGGGGGTCAAGGAGGCTCAGGCCCTGGACTTTGATGTG 2099
QY 2104 TCCACACCATCTACTGACAGAGCTGACCTGAGACCATCAGCGCGGCTTTCATG 2163
DB 2100 TCCACACCATCTACTGACAGAGCTGACCTGAGACCATCAGCGCGGCTTTCATG 2159

QY 2164 AACGGAGCTCGGTGGAGCAGCTGTGTGAGTTTGGCTTGACTACCCGAGGGCATGGCC 2223
DB 2160 AACGGAGCTCGGTGGAGCAGCTGTGTGAGTTTGGCTTGACTACCCGAGGGCATGGCC 2219
QY 2224 GTTGACTGAGTGGGCAAGAACCTTCTACTGGGCCGACACTGGGACCAACAGAAATCGAAATG 2283
DB 2220 GTTGACTGAGTGGGCAAGAACCTTCTACTGGGCCGACACTGGGACCAACAGAAATCGAAATG 2279
QY 2284 CGCGGCTGGACGCGGCTAGTTCCGCAAGTCTCTGTGTGAGAGGACTTGGACAAACCCGAGG 2343
DB 2280 CGCGGCTGGACGCGGCTAGTTCCGCAAGTCTCTGTGTGAGAGGACTTGGACAAACCCGAGG 2339
QY 2344 TCGTGTGGCCCTGGATCCCAACAGGGCTACACTACTGAGACGAGTGGGGCGGCAAGCG 2403
DB 2340 TCGTGTGGCCCTGGATCCCAACAGGGCTACACTACTGAGACGAGTGGGGCGGCAAGCG 2399
QY 2404 AGGATCTGTGGCGGCTTTCATGAGCGGGACCAACTGCACTGAGCGCTGGTGGACAAAGTGGGC 2463
DB 2400 AGGATCTGTGGCGGCTTTCATGAGCGGGACCAACTGCACTGAGCGCTGGTGGACAAAGTGGGC 2459
QY 2464 CGGGCCAAAGACCTTCAACATTTGACTAGCGTGAACGAGCGCTCTACTGAGACGACTGGAC 2523
DB 2460 CGGGCCAAAGACCTTCAACATTTGACTAGCGTGAACGAGCGCTCTACTGAGACGACTGGAC 2519
QY 2524 ACCAATGATCGAGTCTGCAACATGCTGGGTGAGGAGCGGTCTGTGATTCGCGACGAT 2583
DB 2520 ACCAATGATCGAGTCTGCAACATGCTGGGTGAGGAGCGGTCTGTGATTCGCGACGAT 2579
QY 2584 CTCGCGCACCCGTTCTGGTCTGACGAGTACAGCGATTAATATCTACTGAGACAGCTGGAAT 2643
DB 2580 CTCGCGCACCCGTTCTGGTCTGACGAGTACAGCGATTAATATCTACTGAGACAGCTGGAAT 2639
QY 2644 CTGACAGCATTTGAGCGGGCGGACAAAGATAGCGGGCGGAAACCGACCCCTCATCAGGGC 2703
DB 2640 CTGACAGCATTTGAGCGGGCGGACAAAGATAGCGGGCGGAAACCGACCCCTCATCAGGGC 2699
QY 2704 CACCTGAGCTTCTGTGATGAGCATCTCTGTGTTTCACTCTCCCGCAGGATGGCTCAAT 2763
DB 2700 CACCTGAGCTTCTGTGATGAGCATCTCTGTGTTTCACTCTCCCGCAGGATGGCTCAAT 2759
QY 2764 GACTGTATGACAAACCGGCGAGTGTGGCAGCTGTGCTTGGCATCTCCCGGGGGCCAC 2823
DB 2760 GACTGTATGACAAACCGGCGAGTGTGGCAGCTGTGCTTGGCATCTCCCGGGGGCCAC 2819
QY 2824 CGCTGCGGCTGCGGCTCAGACTACACCTTGGACCCGAGCGGACCACTGAGCGCGGCC 2883
DB 2820 CGCTGCGGCTGCGGCTCAGACTACACCTTGGACCCGAGCGGACCACTGAGCGCGGCC 2879
QY 2884 ACCACCTTCTGTGTTTTCAGCAAGAAATCTGCAATGAGTGGATGATCCCGGAGACGAG 2943
DB 2880 ACCACCTTCTGTGTTTTCAGCAAGAAATCTGCAATGAGTGGATGATCCCGGAGACGAG 2939
QY 2944 CACAGCCGGATCTCATCTCGCCCTGCACTGGAATGAGGAACTGCAAGCCATCGACTAT 3003
DB 2940 CACAGCCGGATCTCATCTCGCCCTGCACTGGAATGAGGAACTGCAAGCCATCGACTAT 2999
QY 3004 GACCCACTGACAAAGTTTCATCTACTGAGTGGAGCGGCGGACCAAGATCAGCGAGCCAG 3063
DB 3000 GACCCACTGACAAAGTTTCATCTACTGAGTGGAGCGGCGGCGGACCAAGATCAGCGAGCCAG 3059
QY 3064 GACGAGCGGACCCAGCCCTTTTGTGACCTCTGAGCGCAAGCCCAAAACCCAGACAGG 3123
DB 3060 GACGAGCGGACCCAGCCCTTTTGTGACCTCTGAGCGCAAGCCCAAAACCCAGACAGG 3119
QY 3124 CAGCCCAACGACTCAGCATCTCAGCGGACCTGATGAGCGGACCTGTTTCTGGAAGTGGAGGCC 3183
DB 3120 CAGCCCAACGACTCAGCATCTCAGCGGACCTGATGAGCGGACCTGTTTCTGGAAGTGGAGGCC 3179
QY 3184 ACCAATACCATCAAGTTCAGAGCTGAGCGGGAGCCATGAGGGGTGGTCTGCTGGTGG 3243
DB 3180 ACCAATACCATCAAGTTCAGAGCTGAGCGGGAGCCATGAGGGGTGGTCTGCTGGTGG 3239
QY 3244 GACCGGACCAAGCCCGAGGGCCATCTGCTCAACGCGGAGCGAGGGTACCTGTTACTTACC 3303

Db 3240 GACCGCGAAGCCAGCGCCATCGTCTCAAGCGGAGCGAGGTACTCTGTACTTACC 3299
QY 3304 AACATGAGGACCGGGAGCCCAAGATCGAAGCGCGAGCCCTGGAGCGGACCGAGCGGAG 3363
Db 3300 AACATGAGGACCGGGAGCCCAAGATCGAAGCGCGAGCCCTGGAGCGGACCGAGCGGAG 3359
QY 3364 GTCCTCTTCAACACCGGCTCATCGCCCTGTGGCCCTGTGGCTGTGGAGCAACACACTGGGC 3423
Db 3360 GTCCTCTTCAACACCGGCTCATCGCCCTGTGGCCCTGTGGCTGTGGAGCAACACACTGGGC 3419
QY 3424 AAGCTGTTCTGGTGGAGCGCGAAGCTGAAGCGCATTTGAGAGCTGTGACCTGTCAAGGGCC 3483
Db 3420 AAGCTGTTCTGGTGGAGCGCGAAGCTGAAGCGCATTTGAGAGCTGTGACCTGTCAAGGGCC 3479
QY 3484 AACCGCTGACCTGGAGAGCGCAACATCGTGAGCCCTGTGGCCCTGTGGAGCAACACACTGGGC 3543
Db 3480 AACCGCTGACCTGGAGAGCGCAACATCGTGAGCCCTGTGGCCCTGTGGAGCAACACACTGGGC 3539
QY 3544 AAGCATCTTACTGATCGACCGCCAGCAGCAGATGATCGAGCGTGTGGAGCAACACACTGGGC 3603
Db 3540 AAGCATCTTACTGATCGACCGCCAGCAGCAGATGATCGAGCGTGTGGAGCAACACACTGGGC 3599
QY 3604 GGGACAAAGCGACTCGCATCCAGGGCGTGTGGCCCACTCTGAGGATCCATGCAAGTG 3663
Db 3600 GGGACAAAGCGACTCGCATCCAGGGCGTGTGGCCCACTCTGAGGATCCATGCAAGTG 3659
QY 3664 GAGGAGTCAAGCTGGAGGAGTCTCAGCCCAACCATGTGCGGCTGAGCAATGGTGGCTGC 3723
Db 3660 GAGGAGTCAAGCTGGAGGAGTCTCAGCCCAACCATGTGCGGCTGAGCAATGGTGGCTGC 3719
QY 3724 TCCACATCTGTATTGCAAGGCTGATGGGACACCAAGCTGCTCATGCCAGTCCACCTC 3783
Db 3720 TCCACATCTGTATTGCAAGGCTGATGGGACACCAAGCTGCTCATGCCAGTCCACCTC 3779
QY 3784 GTGCTCTGCAAGAACTGCTGACCTGTGGAGAGCGCCCACTGCTCCCGGACCAAGTTT 3843
Db 3780 GTGCTCTGCAAGAACTGCTGACCTGTGGAGAGCGCCCACTGCTCCCGGACCAAGTTT 3839
QY 3844 GCATGTGCCACAGGGAGATCGACTGTATCCCGGGCTGCGCTGTGAGCGCTTCCC 3903
Db 3840 GCATGTGCCACAGGGAGATCGACTGTATCCCGGGCTGCGCTGTGAGCGCTTCCC 3899
QY 3904 GAGTGCATGACAGAGCGAGGAGGCTGCGCTGCTCCCGCCAGTTCCTCC 3963
Db 3900 GAGTGCATGACAGAGCGAGGAGGCTGCGCTGCTCCCGCCAGTTCCTCC 3959
QY 3964 TGCAGCGGGGTCAAGTGTGGAAGCTGCGCTGCGCTGCGAGCGGAGGAGTGTGAG 4023
Db 3960 TGCAGCGGGGTCAAGTGTGGAAGCTGCGCTGCGCTGCGAGCGGAGGAGTGTGAG 4019
QY 4024 GACCGCTCAGAGGGGAGTGTGAGCGCATCTGCTGCCCAACAGTTCGGTGTGCG 4083
Db 4020 GACCGCTCAGAGGGGAGTGTGAGCGCATCTGCTGCCCAACAGTTCGGTGTGCG 4079
QY 4084 AGCGCGAGTGTGCTCATCAAGCAGAGTGTGAGTCTTCCCGAGTGTGAGCGG 4143
Db 4080 AGCGCGAGTGTGCTCATCAAGCAGAGTGTGAGTCTTCCCGAGTGTGAGCGG 4138
QY 4144 TCCGACGAGTCTCATGTGTGAATCAACAGCGGCTTCCAGACGAGCGGCGCCACAGC 4203
Db 4139 TCCGACGAGTCTCATGTGTGAATCAACAGCGGCTTCCAGACGAGCGGCGCCACAGC 4198
QY 4204 AGTGCATCGGGCGCTCATTTGGCATCATCTCTCTCTCTCTCTCTCTCTCTCTCT 4263
Db 4199 AGTGCATCGGGCGCTCATTTGGCATCATCTCTCTCTCTCTCTCTCTCTCTCTCT 4258
QY 4264 TTTGTGCCAGCGGCTGTGCGAGCGTGTGCGGGGCAAGCGGCGCTTCCCGCAC 4323
Db 4259 TTTGTGCCAGCGGCTGTGCGAGCGTGTGCGGGGCAAGCGGCGCTTCCCGCAC 4318
QY 4324 GAGTATGTGAGGGGAGCCCGCAGCGTGCCTCTCAATTTTCATAGCCCGGGGCTTCCCG 4383

Db 4319 GAGTATGTGAGGGGAGCCCGCAGCGTGCCTCTCAATTTTCATAGCCCGGGGCTTCCCG 4378
QY 4384 CATGGCCCTTTCAGAGCATCGCATCGGAAAGTTCATGATGAGTCCGTGAGCGCTGATG 4443
Db 4379 CATGGCCCTTTCAGAGCATCGCATCGGAAAGTTCATGATGAGTCCGTGAGCGCTGATG 4438
QY 4444 GGGGGCGGGGCGGGGCTCTGTAGACCGGAAACACAGTTCAGAGGGGCTTCTGTCAGC 4503
Db 4439 GGGGGCGGGGCGGGGCTCTGTAGACCGGAAACACAGTTCAGAGGGGCTTCTGTCAGC 4498
QY 4504 AGTCTGTCAGACGAAAGGCGACGCTGTATCCCGCGATTCCTGAAACCGCGCCCTTCCCG 4563
Db 4499 AGTCTGTCAGACGAAAGGCGACGCTGTATCCCGCGATTCCTGAAACCGCGCCCTTCCCG 4558
QY 4564 GCCAGGACCCCTCCCTGTATCAACATGAGCATGTTCTACTCTTCAAAATTCGCGGCACT 4623
Db 4559 GCCAGGACCCCTCCCTGTATCAACATGAGCATGTTCTACTCTTCAAAATTCGCGGCACT 4618
QY 4624 GCGAGACCGTACAGGCGCTTACATCATTCGAGGAATGGCGCCCGGACGACGCGCTTGCAGC 4683
Db 4619 GTGAGACCGTACAGGCGCTTACATCATTCGAGGAATGGCGCCCGGACGACGCGCTTGCAGC 4678
QY 4684 ACCGACGTGTGACAGGAGTACAGCGCAGCGCTGGAAAGGCGAGCAAGTACTTACCTG 4743
Db 4679 ACCGACGTGTGACAGGAGTACAGCGCAGCGCTGGAAAGGCGAGCAAGTACTTACCTG 4738
QY 4744 GATTGAACTCGGACTCAGACCCCTATCCACCCCGGACGACGCGGCGGAGTACCTG 4803
Db 4739 GATTGAACTCGGACTCAGACCCCTATCCACCCCGGACGACGCGGCGGAGTACCTG 4798
QY 4804 TCGGCGGAGGACAGCTCGCGCCCTCGCGCCCGGACCGAGAGGAGTACTTCCATCTCTTC 4863
Db 4799 TCGGCGGAGGACAGCTCGCGCCCTCGCGCCCGGACCGAGAGGAGTACTTCCATCTCTTC 4858
QY 4864 CCGCCCTCTCGTCCCTCGGACGGAATCTATCTGAGCTCGCGCGGCGGAGTCTGCGCTTC 4923
Db 4859 CCGCCCTCTCGTCCCTCGGACGGAATCTATCTGAGCTCGCGCGGCGGAGTCTGCGCTTC 4918
QY 4924 TCTGTGCGGCTGTAAATAGTATTAATATGAACAAAGAAATATATATATATATATAT 4983
Db 4919 TCTGTGCGGCTGTAAATAGTATTAATATGAACAAAGAAATATATATATATATATAT 4978
QY 4984 AAAATAAATATATATGGGATTTTAAAAAATATGAAGAAATGTGAATGTGATGGGTGGC 5043
Db 4979 AAAATAAATATATGGGATTTTAAAAAATATGAAGAAATGTGAATGTGATGGGTGGC 5038
QY 5044 AGGCTCGGAGAACTTTGTA 5063
Db 5039 AGGCTCGGAGAACTTTGTA 5058

RESULT 8

US-09-060-299-32
; Sequence 32, Application US/09060299
; Patent No. 6545137
; GENERAL INFORMATION:
; APPLICANT: Todd, John A
; APPLICANT: Hess, John W
; APPLICANT: Caskey, Charles T
; APPLICANT: Cox, Roger D
; APPLICANT: Gerhold, David
; APPLICANT: Hammond, Holly
; APPLICANT: Hey, Patricia
; APPLICANT: Kawaguchi, Yoshihiko
; APPLICANT: Merriman, Tony R
; APPLICANT: Metzker, Michael L
; TITLE OF INVENTION: No. 6545137el Receptor
; NUMBER OF SEQUENCES: 455
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon and Vanderhye
; STREET: 1100 No. 6545137th Glebe Road, Eighth Floor
; CITY: Arlington
; STATE: Virginia

COUNTRY: US
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/060,299
FILING DATE: 15-APR-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/043,553
FILING DATE: 15-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/048,740
FILING DATE: 05-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: B.J.Sadoff
REGISTRATION NUMBER: 36,663
REFERENCE/DOCKET NUMBER: 620-35
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)816-4091
TELEFAX: (703)816-4100
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 5022 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-060-299-32

Query Match 96.8%; Score 4902; DB 4; Length 5022;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 4916; Conservative 0; Mismatches 5; Indels 1; Gaps 1;
QY 142 GCCTCGCGCTCTCTGCTATTTTCCAAACCGCGGACGCTAGCGCTGGTGGAGCCCGCGCGGA 201
DB 62 GCCTCGCGCTCTCTGCTATTTTCCAAACCGCGGACGCTAGCGCTGGTGGAGCCCGCGCGGA 121
QY 202 GTCAGCTGGAGTCACCATCGTGTGTCAGCGCCCTGGAGGATCGCGCGCGAGTGTTC 261
DB 122 GTCAAGCTGGAGTCACCATCGTGTGTCAGCGCCCTGGAGGATCGCGCGCGAGTGTTC 181
QY 262 CAGTTTTCAGGAGGCGCTGTACTGGACAGACGTGACGAGGAGGCCATCAAGCAGACC 321
DB 182 CAGTTTTCAGGAGGCGCTGTACTGGACAGACGTGACGAGGAGGCCATCAAGCAGACC 240
QY 322 TACCTGAAACAGACGCGGCGCGCTGTGAGAACTGTGTCATCTCGCGCTGTCTCTCCC 381
DB 241 TACCTGAAACAGACGCGGCGCGCTGTGAGAACTGTGTCATCTCGCGCTGTCTCTCCC 300
QY 382 GACGCGCTCGCTGCTGCTGGTGGGCGGAGAGTGTGTCAGCGGACTCAGAGACCAAC 441
DB 301 GACGCGCTCGCTGCTGCTGGTGGGCGGAGAGTGTGTCAGCGGACTCAGAGACCAAC 360
QY 442 CGCATCGAGGTGGCCAACTCAATGGCACATCCCGGAGGTGCTCTCTGGCAGGACCTT 501
DB 361 CGCATCGAGGTGGCCAACTCAATGGCACATCCCGGAGGTGCTCTCTGGCAGGACCTT 420
QY 502 GACCACTAGGCGCATCGCTTGGACCCCGCTCAGCGGTACATGTACTGGACAGACTGG 561
DB 421 GACCACTAGGCGCATCGCTTGGACCCCGCTCAGCGGTACATGTACTGGACAGACTGG 480
QY 562 GGTGAGAGCCCGGATTTGAGCGGCGAGGATGGATGGACACCCCGGAGATCATTTGTG 621
DB 481 GGTGAGAGCCCGGATTTGAGCGGCGAGGATGGATGGACACCCCGGAGATCATTTGTG 540
QY 622 GACTCGGACATTTACTGGCCCAATGGACTGACCATCGACCTGGAGGAGCAGAGCTCTAC 681
DB 541 GACTCGGACATTTACTGGCCCAATGGACTGACCATCGACCTGGAGGAGCAGAGCTCTAC 600
QY 682 TGGGCTGACGCGCAAGCTCAGCTTTCATCCACCGTGCACAACTGGACGGCTCTCGGCGAG 741

DB 601 TGGGCTGACGCAAGCTCAGCTTCATCCACCGTGCACAACTGGACGGCTGTTCCGGCAG 660
QY 742 AAGGTGGTGGAGGAGCCTTCACGACCCCTTCGCGCTGAGCGCTCTCCGGGACACTCTG 801
DB 661 AAGGTGGTGGAGGAGCCTTCACGACCCCTTCGCGCTGAGCGCTCTCCGGGACACTCTG 720
QY 802 TACTGGACAGACTGGCAGACCGCTTCATCGCTGCAACAGCGCACTGGGGGGAAG 861
DB 721 TACTGGACAGACTGGCAGACCGCTTCATCGCTGCAACAGCGCACTGGGGGGAAG 780
QY 862 AGGAAGAGATCCTGAGTGCCTCTTACTCACCCATGGACATCCAGGTCTGAGCCAGAG 921
DB 781 AGGAAGAGATCCTGAGTGCCTCTTACTCACCCATGGACATCCAGGTCTGAGCCAGAG 840
QY 922 CGGCAGCCTTTTCCACACTCGCTGTGAGAGGACAAATGCGGCTGTCTCCACTGTGC 981
DB 841 CGGCAGCCTTTTCCACACTCGCTGTGAGAGGACAAATGCGGCTGTCTCCACTGTGC 900
QY 982 CTGCTGCCCAAGGAGCCTTTCTACATGCGGCTGCCCAAGGCTGTGAGCTGCAG 1041
DB 901 CTGCTGCCCAAGGAGCCTTTCTACACATGCGCTGCCCAAGGCTGTGAGCTGCAG 960
QY 1042 GACAAACGCGACGCTGTAAAGCGAGGAGCCGAGAGGTGCTGCTGCGCCCGCGGACG 1101
DB 961 GACAAACGCGACGCTGTAAAGCGAGGAGCCGAGAGGTGCTGCTGCGCCCGCGGACG 1020
QY 1102 GACCTAGGAGATCTGCTGGACACGCGCGGACTTCCACCGACATGCTGCTGAGGTGAC 1161
DB 1021 GACCTAGGAGATCTGCTGGACACGCGCGGACTTTCACCGACATGCTGCTGAGGTGAC 1080
QY 1162 GACATCCGCGACGCTATTTGCCATCGACTACGACCCGCTAGAGGGCTATGCTACTGGACA 1221
DB 1081 GACATCCGCGACGCTATTTGCCATCGACTACGACCCGCTAGAGGGCTATGCTACTGGACA 1140
QY 1222 GATGACGAGGTGCGGGCCATCCGCGAGGCGTACTGACGCGGTCTGGGGCGCAGACGCTG 1281
DB 1141 GATGACGAGGTGCGGGCCATCCGCGAGGCGTACTGACGCGGTCTGGGGCGCAGACGCTG 1200
QY 1282 GTCAACCGGAGATCAACGAGCCCGATGCGGTGCGATCGCGGTGCGTGGTGGCCCGAAACCTC 1341
DB 1201 GTCAACCGGAGATCAACGAGCCCGATGCGGTGCGATCGCGGTGCGTGGTGGCCCGAAACCTC 1260
QY 1342 TACTGGACCGACGCGGACGCGATCGAGGTGACGCGCTCAAGCGCACCTCCCGC 1401
DB 1261 TACTGGACCGACGCGGACGCGATCGAGGTGACGCGCTCAAGCGCACCTCCCGC 1320
QY 1402 AAGATCTGTGTGTCGAGGACCTGGACGAGCCCGAGCCATGCGACATGCACTGCACTGGATG 1461
DB 1321 AAGATCTGTGTGTCGAGGACCTGGACGAGCCCGAGCCATGCGACATGCACTGCACTGGATG 1380
QY 1462 GGCCTCATGTGTCGACAGACTGGGAGAGAACCTTAATTCGAGTGTGCGCACTTGGAT 1521
DB 1381 GGCCTCATGTGTCGACAGACTGGGAGAGAACCTTAATTCGAGTGTGCGCACTTGGAT 1440
QY 1522 GGGCAGGAGCGGCTGTGCTGCTCAATGCTCTCCCTCGGGTGGCCCAAGCGCTGCGCCCTG 1581
DB 1441 GGGCAGGAGCGGCTGTGCTGCTCAATGCTCTCCCTCGGGTGGCCCAAGCGCTGCGCCCTG 1500
QY 1582 GACTGAGGAGGGAAGTCTTACTGGGAGAGCGCCAAAGACAGAAAGATCAGAGTGTATC 1641
DB 1501 GACTGAGGAGGGAAGTCTTACTGGGAGAGCGCCAAAGACAGAAAGATCAGAGTGTATC 1560
QY 1642 AATGTTGATGGGAGGAGGCGGACCTCTGAGGAGCAGCTCCCGCACTTTTGGG 1701
DB 1561 AATGTTGATGGGAGGAGGCGGACCTCTGAGGAGCAGCTCCCGCACTTTTGGG 1620
QY 1702 TTCACTGCTGTGGGGGACTTTCATCTACTGAGTGTGTCGAGCGCCCGCAGCATCGAGCGG 1761
DB 1621 TTCACTGCTGTGGGGGACTTTCATCTACTGAGTGTGTCGAGCGCCCGCAGCATCGAGCGG 1680
QY 1762 GTGCAAGGTCAAGGCGAGCGGAGCGTCAATTCATGACAGCTGCGCGAGCTGATGGG 1821
DB 1681 GTGCAAGGTCAAGGCGAGCGGAGCGTCAATTCATGACAGCTGCGCGAGCTGATGGG 1740

QY 322 TACCTGAACCAAGACGGGGGCGCCGTGACAGAACGTGGTCACTCCGGGCTGGTCTCTCCC 381
DB |||||
QY 241 TACCTGAACCAAGACGGGGGCGCCGTGACAGAACGTGGTCACTCCGGGCTGGTCTCTCCC 300
DB |||||
QY 382 GACGGCTCGGCTGCGACTGGGTGGGCAAGAGCTGTACTGGAAGGACTGACAGACCAAC 441
DB |||||
QY 301 GACGGCTCGGCTGCGACTGGGTGGGCAAGAGCTGTACTGGAAGGACTGACAGACCAAC 360
DB |||||
QY 442 CGCATCGAGGTGGCCAACTCAATGGCACATCCCGAAGGTGCTCTTCTGGCAGACCTT 501
DB |||||
QY 361 CGCATCGAGGTGGCCAACTCAATGGCACATCCCGAAGGTGCTCTTCTGGCAGACCTT 420
DB |||||
QY 502 GACCAGCCTAGGGCCCATCGCCTTGGACCCCGCTCACGGGTACATGTACTGGAAGAGCTGG 561
DB |||||
QY 421 GACCAGCCGAGGCCCATCGCCTTGGACCCCGCTCACGGGTACATGTACTGGAAGAGCTGG 480
DB |||||
QY 562 GGTGAGACCCCGGATTTGAGCGGCGAGGGATGGATGGCAGACCCCGAAGATCATTTGTG 621
DB |||||
QY 481 GGTGAGACCCCGGATTTGAGCGGCGAGGGATGGATGGCAGACCCCGAAGATCATTTGTG 540
DB |||||
QY 622 GACTCGGACATTTACTGSCCCCAATGGACTGACCATCGACCTGGAGGACAGAGCTCTAC 681
DB |||||
QY 541 GACTCGGACATTTACTGSCCCCAATGGACTGACCATCGACCTGGAGGACAGAGCTCTAC 600
DB |||||
QY 682 TGGGCTGACGCGCAAGCTCAGCTTTCATCCACCGTGCACCTGGACGGCTCGTTCCGGCAG 741
DB |||||
QY 601 TGGGCTGACGCGCAAGCTCAGCTTTCATCCACCGTGCACCTGGACGGCTCGTTCCGGCAG 660
DB |||||
QY 742 AAGGTGTGGAGGGGAGGCTGACGCGACCCCTTGGCCCTTGACGCTCTCCGGGAGCACTCTG 801
DB |||||
QY 661 AAGGTGTGGAGGGGAGGCTGACGCGACCCCTTGGCCCTTGACGCTCTCCGGGAGCACTCTG 720
DB |||||
QY 802 TACTGGAAGAGATCTGAGTGGCCCTTACTCACCCATGGACATCCAGGTGCTGAGCCAGGAG 921
DB |||||
QY 781 AGGAAGAGATCTGAGTGGCCCTTACTCACCCATGGACATCCAGGTGCTGAGCCAGGAG 840
DB |||||
QY 922 CGGAGACCTTTCTTCCACACTGCTGTGAGGAGGACAAATGGCGGTGCTCCACACTGTGTC 981
DB |||||
QY 841 CGGAGACCTTTCTTCCACACTGCTGTGAGGAGGACAAATGGCGGTGCTCCACACTGTGTC 900
DB |||||
QY 982 CTGCTGTCCCAAGCGAGCCTTTCTACACATGCGCCTGCCCAAGGGTGTGACGCTCCAG 1041
DB |||||
QY 901 CTGCTGTCCCAAGCGAGCCTTTCTACACATGCGCCTGCCCAAGGGTGTGACGCTCCAG 960
DB |||||
QY 1042 GACAAACCGCAGGACGTGTAAAGGCAAGGAGCGGAGGAGTGTGCTGCTGGCCCGCGGACG 1101
DB |||||
QY 961 GACAAACCGCAGGACGTGTAAAGGCAAGGAGCGGAGGAGTGTGCTGCTGGCCCGCGGACG 1020
DB |||||
QY 1102 GACCTACGAGGATCTCGCTGGAACACGCGGACATTCACCGACATGCTGCTGCAAGGTGAC 1161
DB |||||
QY 1021 GACCTACGAGGATCTCGCTGGAACACGCGGACATTCACCGACATGCTGCTGCAAGGTGAC 1080
DB |||||
QY 1162 GACATCGGCGACGCAATTTGCCATCGACTACGACCGCTAGAGGGCTATGTCTACTGGACA 1221
DB |||||
QY 1081 GACATCGGCGACGCAATTTGCCATCGACTACGACCGCTAGAGGGCTATGTCTACTGGACA 1140
DB |||||
QY 1222 GATGACGAGGTGCGGGCCCATCCGACGGGCGTACCTGGACCGGTCTGGGGCGCAGACGCTG 1281
DB |||||
QY 1141 GATGACGAGGTGCGGGCCCATCCGACGGGCGTACCTGGACCGGTCTGGGGCGCAGACGCTG 1200
DB |||||
QY 1282 GTCAAACCGGAGATCAACGACCCCGATGGCATTCGCGTGTGACTGGGTGGCCCGAAACCTC 1341
DB |||||
QY 1201 GTCAAACCGGAGATCAACGACCCCGATGGCATTCGCGTGTGACTGGGTGGCCCGAAACCTC 1260
DB |||||
QY 1342 TACTGGAACCGGACCGGACCGGATCGAGGTGACGCGCCTCAACGGGACCTTCCCGC 1401
DB |||||
QY 1261 TACTGGAACCGGACCGGACCGGATCGAGGTGACGCGCCTCAACGGGACCTTCCCGC 1320
DB |||||

QY 1402 AAGATCCTGTGTGTCGAGGACCTTGGAAGAGCCCGAGCCATCGCACTGCACCCCGTGTG 1461
DB |||||
QY 1321 AAGATCCTGTGTGTCGAGGACCTTGGAAGAGCCCGAGCCATCGCACTGCACCCCGTGTG 1380
DB |||||
QY 1462 GGCCTCATGTACTGGAAGACTGGGGAGAGAACCTTAAATCGAGTGTGCCAACTTTGGAT 1521
DB |||||
QY 1381 GGCCTCATGTACTGGAAGACTGGGGAGAGAACCTTAAATCGAGTGTGCCAACTTTGGAT 1440
DB |||||
QY 1522 GGGCAGAGAGCGGCGTGTGCTCAATGCTCCTCGGGTGGCCCAAGCGGCTGCCCTG 1581
DB |||||
QY 1441 GGGCAGAGAGCGGCGTGTGCTCAATGCTCCTCGGGTGGCCCAAGCGGCTGCCCTG 1500
DB |||||
QY 1582 GACCTGAGGAGGGGAAAGCTTACTTGGGGAGACGCAAGACAGACAGATCGAGGTGATC 1641
DB |||||
QY 1501 GACCTGAGGAGGGGAAAGCTTACTTGGGGAGACGCAAGACAGACAGATCGAGGTGATC 1560
DB |||||
QY 1642 AATGTTGATGGGACGAAGAGCGGACCTCTCTGAGGGAAGAAGCTCCCGACATTTTCGG 1701
DB |||||
QY 1561 AATGTTGATGGGACGAAGAGCGGACCTCTCTGAGGGAAGAAGCTCCCGACATTTTCGG 1620
DB |||||
QY 1702 TTCAAGCTGTGCGGGGACTTTCATCTTCTGGAAGCTGACTGGCAGCGCGCAGCATCGAGCG 1761
DB |||||
QY 1621 TTCAAGCTGTGCGGGGACTTTCATCTTCTGGAAGCTGACTGGCAGCGCGCAGCATCGAGCG 1680
DB |||||
QY 1762 GTGCAAGGTCAAGGGCAGCGCGGAGCGTCAATTCATTCACAGCTGCCGACCTGATGG 1821
DB |||||
QY 1681 GTGCAAGGTCAAGGGCAGCGCGGAGCGTCAATTCATTCACAGCTGCCGACCTGATGG 1740
DB |||||
QY 1822 CTCGAAGCTGTGAATGTGGCCAGAGTGTGCGGAACCAACCCGTGTGCGGACAGGAACCG 1881
DB |||||
QY 1741 CTCGAAGCTGTGAATGTGGCCAGAGTGTGCGGAACCAACCCGTGTGCGGACAGGAACCG 1800
DB |||||
QY 1882 GGTGTGAGCACCCTGTGCTTTCACACCCCAAGCTGCTGCGTGGTGGTGGTGGTGGTGGTGG 1941
DB |||||
QY 1801 GGTGTGAGCACCCTGTGCTTTCACACCCCAAGCTGCTGCGTGGTGGTGGTGGTGGTGGTGG 1860
DB |||||
QY 1942 CTGAGCTGTGCTGAGTGACATGAAGACCTGCAATCGTGGCTGAGGGCTTCTTGGTCTTACC 2001
DB |||||
QY 1861 CTGAGCTGTGCTGAGTGACATGAAGACCTGCAATCGTGGCTGAGGGCTTCTTGGTCTTACC 1920
DB |||||
QY 2002 AGCAGAGCGGCATCCACAGGATCTCCCTCGAGACCAATAACAGAGCTGGCCATCCCG 2061
DB |||||
QY 1921 AGCAGAGCGGCATCCACAGGATCTCCCTCGAGACCAATAACAGAGCTGGCCATCCCG 1980
DB |||||
QY 2062 CTCACGGCGCTCAAGAGGCGCTCAGCCCTGAGCTTGTATGTGTCACCAACACACATCTAC 2121
DB |||||
QY 1981 CTCACGGCGCTCAAGAGGCGCTCAGCCCTGAGCTTGTATGTGTCACCAACACACATCTAC 2040
DB |||||
QY 2122 TGGACAGACGTGACCTGAAAGACCATCAGCCGCGCTTTCATGAACGGGAGCTCGGTGGAG 2181
DB |||||
QY 2041 TGGACAGACGTGACCTGAAAGACCATCAGCCGCGCTTTCATGAACGGGAGCTCGGTGGAG 2100
DB |||||
QY 2182 CACGTGTGAGTTTGGCTTGTACTACCCGAGGCGATGGCGGTGACTGATGGGCAAG 2241
DB |||||
QY 2101 CACGTGTGAGTTTGGCTTGTACTACCCGAGGCGATGGCGGTGACTGATGGGCAAG 2160
DB |||||
QY 2242 AACCTCTACTGGGCGGACACTGGGACCAACAGAAATCGAAGTGGCGCGCTGACGGGCGAG 2301
DB |||||
QY 2161 AACCTCTACTGGGCGGACACTGGGACCAACAGAAATCGAAGTGGCGCGCTGACGGGCGAG 2220
DB |||||
QY 2302 TTCCGGCAAGTCTCTGTGTGGAGGGACTTGGACAAACCCGAGGTGCTGGCCCTGATGCC 2361
DB |||||
QY 2221 TTCCGGCAAGTCTCTGTGTGGAGGGACTTGGACAAACCCGAGGTGCTGGCCCTGATGCC 2280
DB |||||
QY 2362 ACCAAGGGCTACATCTTACTGGAACCGAGTGGGGCGGCAAGCCGAGGATCGTGGGCGCTTC 2421
DB |||||
QY 2281 ACCAAGGGCTACATCTTACTGGAACCGAGTGGGGCGGCAAGCCGAGGATCGTGGGCGCTTC 2340
DB |||||
QY 2422 ATGGAACGGGACCAACTGCTGATGACCTGTGGAAGGTGGCGCGGCAAGCCCTCACC 2481
DB |||||
QY 2341 ATGGAACGGGACCAACTGCTGATGACCTGTGGAAGGTGGCGCGGCAAGCCCTCACC 2400
DB |||||
QY 2482 ATTGACTACGTGACCGGCTCTTACTGGACCGGCTGGAACCAACATGATCGAGTGC 2541
DB |||||

|||||
2401 ATTGACTAGCTGACCGAGCCCTCTACTGGACCGACTGGACCAACCAATGATCAGATCG 2460
QY
2542 TCCAAATGCTGGGTGACGAGCGGTGCTGATGTCGCGACGATCTCCGACCGCTTCGGT 2601
Db
2461 TCCAAATGCTGGGTGACGAGCGGTGCTGATGTCGCGACGATCTCCGACCGCTTCGGT 2520
QY
2602 CTGACGCGAGTACAGGATATATCTACTGGACAGATGGAATCTGCACAGCATTTAGCGG 2661
Db
2521 CTGACGCGAGTACAGGATATATCTACTGGACAGATGGAATCTGCACAGCATTTAGCGG 2580
QY
2662 GCCGACAAAGACTAGCGGCGGAAACCGCACCTCTATCCAGGGCACCTGACTTTCGTGATG 2721
Db
2581 GCCGACAAAGACTAGCGGCGGAAACCGCACCTCTATCCAGGGCACCTGACTTTCGTGATG 2640
QY
2722 GACATCTCTGGTGTCCACTCTCTCCGCGCAGGATGCTCAATGACTGTATGCAACAAC 2781
Db
2641 GACATCTCTGGTGTCCACTCTCTCCGCGCAGGATGCTCAATGACTGTATGCAACAAC 2700
QY
2782 GGGCAGTGTGGCAGCTGTGCTTGGCATCCCGGGGCGCACCGCTGGGGCTGCGCCTCA 2841
Db
2701 GGGCAGTGTGGCAGCTGTGCTTGGCATCCCGGGGCGCACCGCTGGGGCTGCGCCTCA 2760
QY
2842 CACTACACCTTGGACCCCGACAGCGCAACTGACCGCGCCACCACTTCTTGTGTTTC 2901
Db
2761 CACTACACCTTGGACCCCGACAGCGCAACTGACCGCGCCACCACTTCTTGTGTTTC 2820
QY
2902 AGCCAGAAATCTGCCATCAGTCTGGATGATCCCGGACGACAGCAGCCCGGATCTCATC 2961
Db
2821 AGCCAGAAATCTGCCATCAGTCTGGATGATCCCGGACGACAGCAGCCCGGATCTCATC 2880
QY
2962 CTGCCCCCTGCACTGAGGAACTCAAGCGAGCCAAAGGAGCGAGCGGACCCAGCGCC 3021
Db
2881 CTGCCCCCTGCACTGAGGAACTCAAGCGAGCCAAAGGAGCGAGCGGACCCAGCGCC 2940
QY
3022 ATCTACTGGGTGGATGGCGCCAGAACTCAAGCGAGCCAAAGGAGCGAGCGGACCCAGCGCC 3081
Db
2941 ATCTACTGGGTGGATGGCGCCAGAACTCAAGCGAGCCAAAGGAGCGAGCGGACCCAGCGCC 3000
QY
3082 TTTGTTTTGACCTCTCTGAGCAGGCGCAAAACCCAGACAGCGAGCCCGACCTCAGC 3141
Db
3001 TTTGTTTTGACCTCTCTGAGCAGGCGCAAAACCCAGACAGCGAGCCCGACCTCAGC 3060
QY
3142 ATCGACATCTACGCGGACACTGTTCTGAGCGTGGAGGCGCCAAATACCATCAAGCTC 3201
Db
3061 ATCGACATCTACGCGGACACTGTTCTGAGCGTGGAGGCGCCAAATACCATCAAGCTC 3120
QY
3202 CACAGGCTGAGCGGGAAAGCCATGGGGTGGTGTGCTGGTGGGAGCCGCAAGCCAGG 3261
Db
3121 CACAGGCTGAGCGGGAAAGCCATGGGGTGGTGTGCTGGTGGGAGCCGCGCAAGCCAGG 3180
QY
3262 GCCATCTGCTCAACCGGAGGAGGTACTGTACTTCAACCAATGACGAGCGGGCA 3321
Db
3181 GCCATCTGCTCAACCGGAGGAGGTACTGTACTTCAACCAATGACGAGCGGGCA 3240
QY
3322 GCCAAGATCGAAACCGGACCGCTTGGACGCGACCGAGCGGAGTCTCTTCAACCAAGCCGCG 3381
Db
3241 GCCAAGATCGAAACCGGACCGCTTGGACGCGACCGAGCGGAGTCTCTTCAACCAAGCCGCG 3300
QY
3382 CTCAATCCGCTCTGGCTTGGTGGTGGACAACTTGGGCAAGCTGTTCTTGGGTGGAC 3441
Db
3301 CTCAATCCGCTCTGGCTTGGTGGTGGACAACTTGGGCAAGCTGTTCTTGGGTGGAC 3360
QY
3442 GCGGACCTGAGCGCATTTAGAGCTGTGACTGTGAGGGGCGAACCGCTGACCTCGGAG 3501
Db
3361 GCGGACCTGAGCGCATTTAGAGCTGTGACTGTGAGGGGCGAACCGCTGACCTCGGAG 3420
QY
3502 GACGCCAATCTGTCAGCTCTGGGCTGACCACTCTTGGCAAGCATCTCTACTGGATC 3561
Db
3421 GACGCCAATCTGTCAGCTCTGGGCTGACCACTCTTGGCAAGCATCTCTACTGGATC 3480
QY
3562 GACGCCAAGCAGATGATCGAGGTGTGAGAGCAACCGGGGCAAGCGGACTCGC 3621
|||||

Db
3481 GACCGCCAGCAGCAGATGATCGAGGCTGTGGAGAACCAACCGGGGCAAGCGGACTCGC 3540
QY
3622 ATCCAGGGCCGTGTGCCCCACTCACTGGCATCCATGAGTGGAGAAAGTCAAGCCTGGAG 3681
Db
3541 ATCCAGGGCCGTGTGCCCCACTCACTGGCATCCATGAGTGGAGAAAGTCAAGCCTGGAG 3600
QY
3682 GAGTTCTCAGGCCCAACCCATGTCGGGTGACAAATGGTGGCTGCTCCACATCTGTATTGCC 3741
Db
3601 GAGTTCTCAGGCCCAACCCATGTCGGGTGACAAATGGTGGCTGCTCCACATCTGTATTGCC 3660
QY
3742 AAGGTGATGGACACACACGCTGTCTATGCCAGTCCACCTCGTCTCTCTGCAGAACCTG 3801
Db
3661 AAGGTGATGGACACACACGCTGTCTATGCCAGTCCACCTCGTCTCTCTGCAGAACCTG 3720
QY
3802 CTGACCTGTGTGAGAGCCGCCCACTGTCTCCCGGACCAAGTTTGCATGTGSCACAGGGGAG 3861
Db
3721 CTGACCTGTGTGAGAGCCGCCCACTGTCTCCCGGACCAAGTTTGCATGTGSCACAGGGGAG 3780
QY
3862 ATCGACTGTATCCCGGGGCTTGGGCTGTGACCGCTTCCGAGTGGATGACAGAGC 3921
Db
3781 ATCGACTGTATCCCGGGGCTTGGGCTGTGACCGCTTCCGAGTGGATGACAGAGC 3840
QY
3922 GACGAGAGGGCTGCCCCGTGTCTCGCGCCAGTTCCTCGCGCGGGTCAAGTGT 3981
Db
3841 GACGAGAGGGCTGCCCCGTGTCTCGCGCCAGTTCCTCGCGCGGGTCAAGTGT 3900
QY
3982 GTGACCTGTGCGCTTGTGCGACGGGAGGACAGACTGTTCAGGACCGGTTCAGAGCGG 4041
Db
3901 GTGACCTGTGCGCTTGTGCGACGGGAGGACAGACTGTTCAGGACCGGTTCAGAGCGG 3960
QY
4042 GACTGTGAGCGCATCTGCTTCCCAACAGTTCGGGTGTGCGGCGGCGGCTGCTCCTC 4101
Db
3961 GACTGTGAGCGCATCTGCTTCCCAACAGTTCGGGTGTGCGGCGGCGGCTGCTCCTC 4020
QY
4102 ATCAAAACAGCAGTGCAGTCTCTTCCCGACTGTATTCGACGGCTCCGACGAGCTCATGTGT 4161
Db
4021 ATCAAAACAGCAGTGCAGTCTCTTCCCGACTGTATTCGACGGCTCCGACGAGCTCATGTGT 4080
QY
4162 GAAATCAGAACCGCGCTTCAGACGACAGCCCGGCGCCACAGAGTGCATCGCGGCCGTC 4221
Db
4081 GAAATCAGAACCGCGCTTCAGACGACAGCCCGGCGCCACAGAGTGCATCGCGGCCGTC 4140
QY
4222 ATTGCGCATCATCTCTCTCTCTGTCATGGGTGTGTCATTTTGTGTCAGCGCGTG 4281
Db
4141 ATTGCGCATCATCTCTCTCTCTGTCATGGGTGTGTCATTTTGTGTCAGCGCGTG 4200
QY
4282 GTGTGCGACGCTATGCGGGGGCCAAACGGGGCTTCCCGCACGAGTATGTTCAGCGGGACC 4341
Db
4201 GTGTGCGACGCTATGCGGGGGCCAAACGGGGCTTCCCGCACGAGTATGTTCAGCGGGACC 4260
QY
4342 CCGCAGCTGCGCTCAATTTTATAGCCCGCGGGCGGTTCCGAGCATGGCCCTTCACAGGC 4401
Db
4261 CCGCAGCTGCGCTCAATTTTATAGCCCGCGGGCGGTTCCGAGCATGGCCCTTCACAGGC 4320
QY
4402 ATCCATGTCGGAAGTCCATGATGAGCTTCGTCAGCTGTATGGGGGGCCGGGGGGTG 4461
Db
4321 ATCCATGTCGGAAGTCCATGATGAGCTTCGTCAGCTGTATGGGGGGCCGGGGGGTG 4380
QY
4462 CCCCTGTACGACCGGAACCAAGTTCAGAGGGGCTTCGTCAGAGAGCTTCGTCAGACGAG 4521
Db
4381 CCCCTGTACGACCGGAACCAAGTTCAGAGGGGCTTCGTCAGAGAGCTTCGTCAGACGAG 4440
QY
4522 GCCAGCTGTACCGCGCATCTTGAAACCGCGCGCTTCCCGGCGACGGACCCCTCCCTG 4581
Db
4441 GCCAGCTGTACCGCGCATCTTGAAACCGCGCGCTTCCCGGCGACGGACCCCTCCCTG 4500
QY
4582 TACAAATGAGCAGTGTCTACTCTTCAAAATTCGCGCCACTGCGAGACCGGTACAGGGCC 4641
Db
4501 TACAAATGAGCAGTGTCTACTCTTCAAAATTCGCGCCACTGCGAGACCGGTACAGGGCC 4560
QY
4642 TACATCATTCGAGGAATGGGCGCCCGCCGACGAGCGCTTCGAGACCGAGCTGTGTGACAGC 4701
Db
4561 TACATCATTCGAGGAATGGGCGCCCGCCGACGAGCGCTTCGAGACCGAGCTGTGTGACAGC 4620

QY 4702 GACTACAGCGCAGCGCTGGAAGGCCAGCAAGTACTACCTGGATTTGAACTCGACTCA 4761
D 4621 GACTACAGCGCAGCGCTGGAAGGCCAGCAAGTACTACCTGGATTTGAACTCGACTCA 4680
QY 4762 GACCCCTATCCACCCCAACCCACCGCCCAACAGCCGACGATGCTGCGCGGAGACAGCTGC 4821
D 4681 GACCCCTATCCACCCCAACCCACCGCCCAACAGCCGACGATGCTGCGCGGAGACAGCTGC 4740
QY 4822 CGCGCTCGCCCGCCAGCAGAGAGCTACTTCCATCTCTTCCCGCCCGCTCCGTCGCC 4881
D 4741 CGCGCTCGCCCGCCAGCAGAGAGCTACTTCCATCTCTTCCCGCCCGCTCCGTCGCC 4800
QY 4882 TGCACGGACTCATCTGACCTCGCGCGGCCACCTCTGCTGCTCTCTGCTGCTCTGTAATA 4941
D 4801 TGCACGGACTCATCTGACCTCGCGCGGCCACCTCTGCTGCTCTCTGCTGCTCTGTAATA 4860
QY 4942 GTTTTAAATATGAACAAAGAAAAATATATTTTATGATTTTAAAAATAATAATATTTG 5001
D 4861 GTTTTAAATATGAACAAAGAAAAATATATTTTATGATTTTAAAAATAATAATATTTG 4920
QY 5002 GATTTTAAAAACATGAGAAATGTGAACCTGATGGGTGGCGAGGCTGGGAGACTTTG 5061
D 4921 GATTTTAAAAACATGAGAAATGTGAACCTGATGGGTGGCGAGGCTGGGAGACTTTG 4980
QY 5062 TA 5063
D 4981 TA 4982

RESULT 10

US-09-060-299-31
; Sequence 31, Application US/09060299
; Patent No. 6545137
; GENERAL INFORMATION:
; APPLICANT: Todd, John A
; APPLICANT: Hess, John W
; APPLICANT: Caskey, Charles T
; APPLICANT: Cox, Roger D
; APPLICANT: Gerhold, David
; APPLICANT: Hammond, Holly
; APPLICANT: Hey, Patricia
; APPLICANT: Kawaguchi, Yoshihiko
; APPLICANT: Merriman, Tony R
; APPLICANT: Metzker, Michael L
; TITLE OF INVENTION: No. 6545137el Receptor
; NUMBER OF SEQUENCES: 455
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon and Vanderhye
; STREET: 1100 No. 6545137th Glebe Road, Eighth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: US
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/060,299
; FILING DATE: 15-APR-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/043,553
; FILING DATE: 15-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/048,740
; FILING DATE: 05-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: B.J.Sadoff
; REGISTRATION NUMBER: 36,663
; REFERENCE/DOCKET NUMBER: 620-35
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703)816-4091
; TELEFAX: (703)816-4100
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5263 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-060-299-31

Query Match 96.7%; Score 4894.4; DB 4; Length 5263;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 4920; Conservative 0; Mismatches 6; Indels 2; Gaps 2;

QY 136 GCCCGCGCTCGCGCTCTCTCTATTTTGCCTAAACCGCGGGACGTACGGCTGGTGACGCC 195
D 298 GCCCGAGCTCGCGCTCTCTCTATTTTGCCTAAACCGCGGGACGTACGGCTGGTGACGCC 357
QY 196 GCGGAGTCAAGCTGGAGTCCACCATCGTGGTCAAGCGGCTGGAGGATCGCGCGCAGTG 255
D 358 GCGGAGTCAAGCTGGAGTCCACCATCGTGGTCAAGCGGCTGGAGGATCGCGCGCAGTG 417
QY 256 GACTTCAGTTTCCAAAGGAGCGCTGTACTGGACAGACGTGAGGAGGACCATCAAG 315
D 418 GACTTCAGTTTCCAAAGGAGCGCTGTACTGGACAGACGTGAGGAGGACCATCAAG 477
QY 316 CAGACCTACCTGAACACAGAGCGGGGCGCGCTGCAGAACGTGCTCATCTCCGGCTGGTC 375
D 478 CAGACCTACCTGAACACAGAGCGGGGCGCGCTGCAGAACGTGCTCATCTCCGGCTGGTC 537
QY 376 TCTCCCGACGGCTCGCTCGACTGGGTGGGCAAGAGCTGTACTTGGACGGACTCAGAG 435
D 538 TCTCCCGACGGCTCGCTCGACTGGGTGGGCAAGAGCTGTACTTGGACGGACTCAGAG 597
QY 436 ACCAACCGCATCGAGGTGGCCAACTCAATGGCAATCCCGGAAGGTGCTCTTCTGGGAG 495
D 598 ACCAACCGCATCGAGGTGGCCAACTCAATGGCAATCCCGGAAGGTGCTCTTCTGGGAG 657
QY 496 GACCTTGACACCGCTAGCGGCTCGCTTGGACCGCGCTCAGCGGTACATGCTACTGGACA 555
D 658 GACCTTGACACCGCGAGGGCCATCGCTTGGACCGCGCTCAGCGGTACATGCTACTGGACA 717
QY 556 GACTGGGGTGAGACCGCCCGGATTTGAGCGGCGAGGGATGGATGGCAGCACCCCGGAAGATC 615
D 718 GACTGGGGTGAGACCGCCCGGATTTGAGCGGCGAGGGATGGATGGCAGCACCCCGGAAGATC 777
QY 616 ATTGTGACTCGGACATTTACTGGCCCAATGGATGACCATCGACTGGAGGAGCAGAG 675
D 778 ATTGTGACTCGGACATTTACTGGCCCAATGGATGACCATCGACTGGAGGAGCAGAG 837
QY 676 CTCTACTGGGCTGACGCGCAAGCTCAGCTTCAATCCACCGTGCACACCTGGACGGCTCGTTC 735
D 838 CTCTACTGGGCTGACGCGCAAGCTCAGCTTCAATCCACCGTGCACACCTGGACGGCTCGTTC 897
QY 736 CGGCAGAGGTGGTGAGGGGAGCTTGAACGACACCCCTTTCGCCCTTTCGCCCTTTCGCCGGGAC 795
D 898 CGGCAGAGGTGGTGAGGGGAGCTTGAACGACACCCCTTTCGCCCTTTCGCCCTTTCGCCGGGAC 957
QY 796 ACTCTGTACTGGACAGACTGGCAGACCGGCTTCATTCATGCTGCAACAGGCGACTGGG 855
D 958 ACTCTGTACTGGACAGACTGGCAGACCGGCTTCATTCATGCTGCAACAGGCGACTGGG 1017
QY 856 GGGAGAGGAGGAGATCTTGAAGTGGCCCTTCTACTCACCCTATGGACATCCAGGTGCTGAGC 915
D 1018 GGGAGAGGAGGAGATCTTGAAGTGGCCCTTCTACTCACCCTATGGACATCCAGGTGCTGAGC 1077
QY 916 CAGGAGCGGAGCGCTTTCTTCCACACTCGCTGTGAGGAGGACAAATGGCGGTGCTCCAC 975
D 1078 CAGGAGCGGAGCGCTTTCTTCCACACTCGCTGTGAGGAGGACAAATGGCGGTGCTCCAC 1137
QY 976 CTGTGCTGTCTTCCCAAGGAGCGCTTTTCAACATGCTGCGCTTCCCGCCAGCGGTGTGAG 1035
D 1138 CTGTGCTGTCTTCCCAAGGAGCGCTTTTCAACATGCTGCGCTTCCCGCCAGCGGTGTGAG 1197

1036 CTGCAGGCAACCGGCAGGACGTGTAAAGCAGAGCCGAGGAGTCTGCTGTCGCCCGG 1095
1198 CTGCAGGCAACCGGCAGGACGTGTAAAGCAGAGCCGAGGAGTCTGCTGTCGCCCGG 1257
1096 CGGACGGACCTACGGAGGATCTCGCTGGACACGCGCGACTTCACCGACATCTGCTGCGAG 1155
1258 CGGACGGACCTACGGAGGATCTCGCTGGACACGCGCGACTTCACCGACATCTGCTGCGAG 1317
1156 GTGGACGACATCCGGCAGCCATTTGCCATCGACTACGACCCCGCTAGAGGGCTATGCTTAC 1215
1318 GTGGACGACATCCGGCAGCCATTTGCCATCGACTACGACCCCGCTAGAGGGCTATGCTTAC 1377
1216 TGGACAGATGACGAGGTGCGGSCCATCCGCGAGGCGTACTGGACGGGTCTGGGCGCGAG 1275
1378 TGGACAGATGACGAGGTGCGGSCCATCCGCGAGGCGTACTGGACGGGTCTGGGCGCGAG 1437
1276 AGCTTGGTCAACACCGAGATCAACACCCCGATCGCGTTCGACTGGGTGCGCCCGA 1335
1438 AGCTTGGTCAACACCGAGATCAACACCCCGATCGCGTTCGACTGGGTGCGCCCGA 1497
1336 AACCTTACTGGACCGACACGCGGACCGGACCGCATCGAGGTGACGCGCTCAACGGCAC 1395
1498 AACCTTACTGGACCGACACGCGGACCGGACCGCATCGAGGTGACGCGCTCAACGGCAC 1557
1396 TCCCGCAAGATCCTGGTGTGCGAGGACCTGGACGAGCCCGAGCCATCGCATCGACCCC 1455
1558 TCCCGCAAGATCCTGGTGTGCGAGGACCTGGACGAGCCCGAGCCATCGCATCGACCCC 1617
1456 GTGATGGGCTCATGTACTGGACAGACTGGGAGAGAACCTTAAATCGAGTGTGCCAAC 1515
1618 GTGATGGGCTCATGTACTGGACAGACTGGGAGAGAACCTTAAATCGAGTGTGCCAAC 1677
1516 TTGGATGGGACGAGCGCGTGTGCTGTCATGCTCTCCCTCGGGTGGCCCAAGCGGCTG 1575
1678 TTGGATGGGACGAGCGCGTGTGCTGTCATGCTCTCCCTCGGGTGGCCCAAGCGGCTG 1737
1576 GCCCTGGACCTGACGAGGAGGAGCTCTACTGGGAGAGCCCAAGACAGACAGATCGAG 1635
1738 GCCCTGGACCTGACGAGGAGGAGCTCTACTGGGAGAGCCCAAGACAGACAGATCGAG 1797
1636 GTGATCAATGTGTGAGCAAGAGGCGGACCTCTCTGGAGGACCAAGCTCCCGCACATT 1695
1798 GTGATCAATGTGTGAGCAAGAGGCGGACCTCTCTGGAGGACCAAGCTCCCGCACATT 1857
1696 TTCGGTTTCACTGCTGGGAGCTTTCATCTACTGGACTGATGCGCAGCGCGCAGCATC 1755
1858 TTCGGTTTCACTGCTGGGAGCTTTCATCTACTGGACTGATGCGCAGCGCGCAGCATC 1917
1756 GAGCGGGTGCAACAGGTCAAGCCAGCCGAGACGTTCATCTTGACCAAGCTGCGGACCTG 1815
1918 GAGCGGGTGCAACAGGTCAAGCCAGCCGAGACGTTCATCTTGACCAAGCTGCGGACCTG 1977
1816 ATGGGGTCAAAAGCTGTGAATGTGGCCAAAGTCTGTCGGAACCAACCCGTGTGGGACAGG 1875
1978 ATGGGGTCAAAAGCTGTGAATGTGGCCAAAGTCTGTCGGAACCAACCCGTGTGGGACAGG 2037
1876 AACGGGGGTGACGACCTGTGCTTTCATCACCACCGCAACCCCGTGTGGCTGCCCC 1935
2038 AACGGGGGGTG-AGCCACTGTGCTTTCATCACCACCGCAACCCCGTGTGGCTGCCCC 2096
1936 ATCGGCTTGAAGTGTGCTGAGTGAATGAAGACCTGTCATGTCGCTGAGGCGCTTCTTGGTC 1995
2097 ATCGGCTTGAAGTGTGCTGAGTGAATGAAGACCTGTCATGTCGCTGAGGCGCTTCTTGGTC 2156
1996 TTCCACGACGAGCGCCATCCACAGGATCTCCCTCGAGACCAATAACACAGCTGGCC 2055
2157 TTCCACGACGAGCGCCATCCACAGGATCTCCCTCGAGACCAATAACACAGCTGGCC 2216
2056 ATCCGCTCAGGGGCTCAAGGAGGCTCAGCGCTGAGCTTTGATGTGTCACCAACCCAC 2115
2217 ATCCGCTCAGGGGCTCAAGGAGGCTCAGCGCTGAGCTTTGATGTGTCACCAACCCAC 2276

2116 ATCTACTGGACAGAGCTCAGCTCAGACCATCAGCGCGCTTCATGAACGGGAGCTCG 2175
2277 ATCTACTGGACAGAGCTCAGCTCAGACCATCAGCGCGCTTCATGAACGGGAGCTCG 2336
2176 GTGGAGCAGCTGTGTGGAGTTTGGGCTTGTACTACCCCGAGGCGATCGCGCTTGTACTGGATG 2235
2337 GTGGAGCAGCTGTGTGGAGTTTGGGCTTGTACTACCCCGAGGCGATCGCGCTTGTACTGGATG 2396
2236 GGCAGAACTCTACTGTGGCGCGACACTGGGACCAACAGATCGAAGTGGCGGCTGAGAC 2295
2397 GGCAGAACTCTACTGTGGCGCGACACTGGGACCAACAGATCGAAGTGGCGGCTGAGAC 2456
2296 GGCAGTTTCCGCGAAGTCTCGTGTGGAGGACTTGGACAAACCCGAGGTGCTGCCCTG 2355
2457 GGCAGTTTCCGCGAAGTCTCGTGTGGAGGACTTGGACAAACCCGAGGTGCTGCCCTG 2516
2356 GATCCCAACAAAGGCTACATCTACTGGACCGAGTGGGCGGCAAGCCGAGGATCGTGGG 2415
2517 GATCCCAACAAAGGCTACATCTACTGGACCGAGTGGGCGGCAAGCCGAGGATCGTGGG 2576
2416 GCCTTCATGGACGGGACCAACTGATGACGCTGGTGGACAAGGTGGGCGGCGCAACGAC 2475
2577 GCCTTCATGGACGGGACCAACTGATGACGCTGGTGGACAAGGTGGGCGGCGCAACGAC 2636
2476 CTCACCATTTGACTACGCTGACCGCTCTACTGGACCGGACCTGGACCAACATGATC 2535
2637 CTCACCATTTGACTACGCTGACCGCTCTACTGGACCGGACCTGGACCAACATGATC 2696
2536 GAGTCGTCCAACATGCTGGGTGAGGAGCGGTCGCTGATGTCGCGACGATCTCCCGCACCG 2595
2697 GAGTCGTCCAACATGCTGGGTGAGGAGCGGTCGCTGATGTCGCGACGATCTCCCGCACCG 2756
2596 TTCCGTCTGACGAGTACAGCGATTATCTACTGGACAGACTGGAATCTGCAACGATT 2655
2757 TTCCGTCTGACGAGTACAGCGATTATCTACTGGACAGACTGGAATCTGCAACGATT 2816
2656 GAGCGGCGCGACAAGACTAGCGGCGGAAACCGCACCTTCATCCAGGGCACCTGAGCTTC 2715
2817 GAGCGGCGCGACAAGACTAGCGGCGGAAACCGCACCTTCATCCAGGGCACCTGAGCTTC 2876
2716 GTGATGACATCTCTGCTTCCACTCTCCCGCAGGATGGCTCAATGACTGTATGCAC 2775
2877 GTGATGACATCTCTGCTTCCACTCTCCCGCAGGATGGCTCAATGACTGTATGCAC 2936
2776 AACAAACGGGCGAGTGTGGGAGCTGTGCTTGCATCCCGCGCGGCGCACCGCTGGGCTGC 2835
2937 AACAAACGGGCGAGTGTGGGAGCTGTGCTTGCATCCCGCGCGGCGCACCGCTGGGCTGC 2996
2836 GCCTTCACACTACACCTTGGACCCCGACGAGCGGCAACTGACGCGCGCGCACCACTTCTTG 2895
2997 GCCTTCACACTACACCTTGGACCCCGACGAGCGGCAACTGACGCGCGCGCACCACTTCTTG 3056
2896 CTGTTTCAGCCAGAAATCTGCCATCAGTGGATGATCCCGGACGACCAAGCACAGCCCGAT 2955
3057 CTGTTTCAGCCAGAAATCTGCCATCAGTGGATGATCCCGGACGACCAAGCACAGCCCGAT 3116
2956 CTGATCTGCGCTGCTGAGTGTGAGGACGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 3015
3117 CTGATCTGCGCTGCTGAGGACTGAGGAACTGAGGAACTGAGGAACTGAGGAACTGAGGAACT 3176
3016 AAGTTTCATCTACTGGGTGGGCGCGCAACATCAAGCGAGGCAAGGACGAGGAGGAC 3075
3177 AAGTTTCATCTACTGGGTGGGCGCGCAACATCAAGCGAGGCAAGGACGAGGAGGAC 3236
3076 CAGCCCTTTGTTTGAATCTCTGAGCCAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCA 3135
3237 CAGCCCTTTGTTTGAATCTCTGAGCCAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCA 3296
3136 CTGAGCATCGACATCTACGCGGACACTGTTCTGAGCTGTCGAGGCGGCAAGGCAAGGCAAGGCA 3195
3297 CTGAGCATCGACATCTACGCGGACACTGTTCTGAGCTGTCGAGGCGGCAAGGCAAGGCAAGGCA 3356
3196 AACGTCCACAGGCTGAGCGGCGAAGCCATGGGGGTGGTGTGCTGCTGGGGAACCGGCAAG 3255

Db 3357 AACGTCACAGGCTGAGCGGGAGCCATGCGGGTGGTGTGCGTGGGACCGGACAA 3416
QY CCCAGGGCCATGTCGTCAACCGCGAGCGAGGTACCTGTACTTCAACCAACATGACGAGC 3315
Db 3417 CCCAGGGCCATGTCGTCAACCGCGAGCGAGGTACCTGTACTTCAACCAACATGACGAGC 3476
QY CGGGCAGCAAGATCGAAGCGGCGAGCCCTCGACGCGACCGGAGCGAGGTCTCTTTCACC 3375
Db 3477 CGGGCAGCAAGATCGAAGCGGCGAGCCCTCGACGCGACCGGAGCGAGGTCTCTTTCACC 3536
QY ACCGGCCTCATCGCCCTGTGGCCCTGTGTGTGGGCAACACATGGGGCAAGCTGTTCGG 3435
Db 3537 ACCGGCCTCATCGCCCTGTGGCCCTGTGTGTGGGCAACACATGGGGCAAGCTGTTCGG 3596
QY GTGGACGGGACCTGAAGCGCATTTGAGAGCTGTGACCTGTGAGGGGCCAACCGGCTGACC 3495
Db 3597 GTGGACGGGACCTGAAGCGCATTTGAGAGCTGTGACCTGTGAGGGGCCAACCGGCTGACC 3656
QY CTGGAGGACGCGCAACATCGTCAGCGCTCTGGGCGCTGACCATCTTGGCAAGCATCTCTAC 3555
Db 3657 CTGGAGGACGCGCAACATCGTCAGCGCTCTGGGCGCTGACCATCTTGGCAAGCATCTCTAC 3716
QY TGGATCGACCGCGACGACGATGATCGAGCGTGTGGAGAGACCAACCGGGCAAGCGG 3615
Db 3717 TGGATCGACCGCGACGACGATGATCGAGCGTGTGGAGAGACCAACCGGGCAAGCGG 3776
QY ACTCGCATCCAGGGCGGTGTCGCCACCTCACTGGCATCCATGCGATGCGAGGAGTCA 3675
Db 3777 ACTCGCATCCAGGGCGGTGTCGCCACCTCACTGGCATCCATGCGATGCGAGGAGTCA 3836
QY CTGGAGGAGTTCTCAGCCCAACCATGTGCGCGTGTGACATGTGGTGTGCTCCACATCTGT 3735
Db 3837 CTGGAGGAGTTCTCAGCCCAACCATGTGCGCGTGTGACATGTGGTGTGCTCCACATCTGT 3896
QY ATTGCCAAGGATGATGGACACCAAGGTGCTCATGCCAGTCCACCTCGTCTCTCTGCGAG 3795
Db 3897 ATTGCCAAGGATGATGGACACCAAGGTGCTCATGCCAGTCCACCTCGTCTCTCTGCGAG 3956
QY AACCTGCTGACTGTGGAGAGCGCCACCTGCTCCCGGACCAAGTTTGCATGTGCCACA 3855
Db 3957 AACCTGCTGACTGTGGAGAGCGCCACCTGCTCCCGGACCAAGTTTGCATGTGCCACA 4016
QY GGGGAGATCGACTGTATCCCGGGGCGCTGGCGCTGTGACGCTTCCCGAGTGGCGATGAC 3915
Db 4017 GGGGAGATCGACTGTATCCCGGGGCGCTGGCGCTGTGACGCTTCCCGAGTGGCGATGAC 4076
QY CAGAGCGACGAGGAGGGCTGCCCGTGTGCTCCGCGCCAGTTCCCTGCGCGCGGGGT 3975
Db 4077 CAGAGCGACGAGGAGGGCTGCCCGTGTGCTCCGCGCCAGTTCCCTGCGCGCGGGGT 4135
QY CAGTGTGTGGACCTGCGCTGCGTGGACGCGGAGGAGGAGTGTGAGGACCGTCA 4035
Db 4136 CAGTGTGTGGACCTGCGCTGCGTGGACGCGGAGGAGGAGTGTGAGGACCGTCA 4195
QY GAGCGGACTGTGACGCGCATCTGCTGCGCCCAACAGTTCCCGTGTGCGAGCGGCGAGTGT 4095
Db 4196 GAGCGGACTGTGACGCGCATCTGCTGCGCCCAACAGTTCCCGTGTGCGAGGCGGCGAGTGT 4255
QY GTCCTCATCAACAGCAGTGGACTCTCTTCCCGGACTGTATCGAGCGGTCCGACGAGTCT 4155
Db 4256 GTCCTCATCAACAGCAGTGGACTCTCTTCCCGGACTGTATCGAGCGGTCCGACGAGTCT 4315
QY ATGTGTGAAATCACCAGCGCCCTCAGACGACGCGCGGCCACAGCAGTGGCCATCGG 4215
Db 4316 ATGTGTGAAATCACCAGCGCCCTCAGACGACGCGCGGCCACAGCAGTGGCCATCGG 4375
QY CCCGTATTGTCATCATCTCTCTCTCTCTGTCATGGGTGTGCTATTTTGTGTCAG 4275
Db 4376 CCCGTATTGTCATCATCTCTCTCTCTCTGTCATGGGTGTGCTATTTTGTGTCAG 4435
QY CGCGTGTGTGCCAGCGTATCGGGGGGCCAAGCGGCGCTTCCCGCAGCAGTATGTCA 4335

Db 4436 CGCGTGTGTGCCAGCGCTATGTCGGGGGCCAAACGGGCGCTTCCGCGACAGTATGTCA 4495
QY GGGACCGCGACGTCGCGCTCAATTTTATAGCCCGGGGGTTCACAGCATGGCCCTTC 4395
Db 4496 GGGACCGCGACGTCGCGCTCAATTTTATAGCCCGGGGGTTCACAGCATGGCCCTTC 4555
QY ACAGGCATCGCATCGGAAAGTCCATGATGAGCTCCGTGAGCCTGATGGGGGGCGGGGC 4455
Db 4556 ACAGGCATCGCATCGGAAAGTCCATGATGAGCTCCGTGAGCCTGATGGGGGGCGGGGC 4615
QY GGGGTGCCCTCTACGACCGGAAACGTCACAGGGGCTCGTCCAGCAGCTCGTCCAGC 4515
Db 4616 GGGGTGCCCTCTACGACCGGAAACGTCACAGGGGCTCGTCCAGCAGCTCGTCCAGC 4675
QY ACAGAGGCGACGCTGTACCGCGCATCTCTGAAACCGCGCGCTCCCGGGCACGAGACCC 4575
Db 4676 ACAGAGGCGACGCTGTACCGCGCATCTCTGAAACCGCGCGCTCCCGGGCACGAGACCC 4735
QY TCCCTGTACAAATGAGCATGTCTTCTTCTTCAACATTCGCGGCACCTGCGAGACCGTAC 4635
Db 4736 TCCCTGTACAAATGAGCATGTCTTCTTCTTCAACATTCGCGGCACCTGCGAGACCGTAC 4795
QY AGGCCCTACATCTCGAGGAATGGCGCCCCCGACGACGCGCTTGACGACCGTGT 4695
Db 4796 AGGCCCTACATCTCGAGGAATGGCGCCCCCGACGACGCGCTTGACGACCGTGT 4855
QY GACAGCGACTACAGCGCGCGCTGGAGGCGGAGTACTACCTGGAATTTGAACCTG 4755
Db 4856 GACAGCGACTACAGCGCGCGCTGGAGGCGGAGTACTACCTGGAATTTGAACCTG 4915
QY GACTCAGACCCCTTACACCCCGCCACCGCCACAGCCAGTACTCTGCGCGGAGGAC 4815
Db 4916 GACTCAGACCCCTTACACCCCGCCACCGCCACAGCCAGTACTCTGCGCGGAGGAC 4975
QY AGCTGCGCGCGCTCGCGCGCCACCGAGAGGAGTACTTCCATCTCTTCCCGCGCCCTCG 4875
Db 4976 AGCTGCGCGCGCTCGCGCGCCACCGAGAGGAGTACTTCCATCTCTTCCCGCGCCCTCG 5035
QY TCCCGCTGCGAGGACTCATCTGACCTCGCGCGGCGCACTCTGCTTCTCTGTGCCCCCTG 4935
Db 5036 TCCCGCTGCGAGGACTCATCTGACCTCGCGCGGCGCACTCTGCTTCTCTGTGCCCCCTG 5095
QY TAAATAGTTTTAAATATGAAACAAAGAAAAATATTTTATGATTTAAAAAATAAATAT 4995
Db 5096 TAAATAGTTTTAAATATGAAACAAAGAAAAATATTTTATGATTTAAAAAATAAATAT 5155
QY AATTGGGATTTAAAAACATGAGAAATGTGAACTGTGATGGGTGGGCGAGGCTGGGAGA 5055
Db 5156 AATTGGGATTTAAAAACATGAGAAATGTGAACTGTGATGGGTGGGCGAGGCTGGGAGA 5215
QY ACTTTGTA 5063
Db 5216 ACTTTGTA 5223

RESULT 11
US-09-923A-31
; Sequence 31, Application US/09402923A
; Patent No. 6555654
; GENERAL INFORMATION:
; APPLICANT: Todd, John A
; Hess, John W
; Caskey, Charles T
; Cox, Roger D
; Gerhold, David
; Hammond, Holly
; Hey, Patricia
; Kawaguchi, Yoshihiko
; Merriman, Tony R
; Metzker, Michael L
; TITLE OF INVENTION: No. 6555654el LDL-Receptor
; NUMBER OF SEQUENCES: 455
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Nixon and Vanderhye
STREET: 1100 No. 655564th Glebe Road, Eighth Floor
CITY: Arlington
STATE: Virginia
COUNTRY: US
ZIP: VA 22201-4714

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25 (BPO)
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/402,923A

FILING DATE: 14-Feb-2001

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/GB98/01102

FILING DATE: 15-APR-1998

APPLICATION NUMBER: US 60/043,553

FILING DATE: 15-APR-1997

APPLICATION NUMBER: US 60/048,740

FILING DATE: 05-JUN-1997

ATTORNEY/AGENT INFORMATION:

NAME: B.J.Sadoff

REGISTRATION NUMBER: 36,663

REFERENCE/DOCKET NUMBER: 620-81

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703)816-4091

TELEFAX: (703)816-4100

INFORMATION FOR SEQ ID NO: 31:

SEQUENCE CHARACTERISTICS:

LENGTH: 5263 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 31:

US-09-402-923A-31

Query Match 96.7%; Score 4894.4; DB 4; Length 5263;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 4920; Conservative 0; Mismatches 6; Indels 2; Gaps 2;

QY	616	ATTGTGGACTCGGACATTTTACTGGCCCAATGGACTGACCATCGACCTCGAGAGCAGAAAG	675
DB	778	ATTGTGGACTCGGACATTTTACTGGCCCAATGGACTGACCATCGACCTCGAGAGCAGAAAG	837
QY	676	CTCTACTGGGCTGACGCCAAGCTCAGCTTATCCACCGTGCACCACTCGAGCGCTCGTTC	735
DB	838	CTCTACTGGGCTGACGCCAAGCTCAGCTTATCCACCGTGCACCACTCGAGCGCTCGTTC	897
QY	736	CGGAGAAAGTGGTGGAGGGCAGCTGACGACCCCTTCGCCCCTGACGCTCTCCGGGGAC	795
DB	898	CGGAGAAAGTGGTGGAGGGCAGCTGACGACCCCTTCGCCCCTGACGCTCTCCGGGGAC	957
QY	796	ACTCTGTACTGGACAGACTGSCAGACCGGCTCCATCCATGCTGCAACAAGCGACTGGG	855
DB	958	ACTCTGTACTGGACAGACTGSCAGACCGGCTCCATCCATGCTGCAACAAGCGACTGGG	1017
QY	856	GGGAGAGAGAGGAGATCCTGAGTGCCCTCTACTCACCCTATGGACATCAGGTGCTGAGC	915
DB	1018	GGGAGAGAGAGGAGATCCTGAGTGCCCTCTACTCACCCTATGGACATCAGGTGCTGAGC	1077
QY	916	CAGAGCGGCGAGCCTTTCTTCCACACTGCTGTGAGGAGACAATGGCGGTGCTGCCAC	975
DB	1078	CAGAGCGGCGAGCCTTTCTTCCACACTGCTGTGAGGAGACAATGGCGGTGCTGCCAC	1137
QY	976	CTGTGCTGCTGCTGCCAAGCGAGCCTTTCTACACATGCGCTGCCCTGCCACCGGCTGTCAG	1035
DB	1138	CTGTGCTGCTGCTGCCAAGCGAGCCTTTCTACACATGCGCTGCCCTGCCACCGGCTGTCAG	1197
QY	1036	CTGAGGACACACGAGCGAGCGTGTAAAGCAGAGCGCGAGGAGGTGCTGCTGCTGCCCGG	1095
DB	1198	CTGAGGACACACGAGCGAGCGTGTAAAGCAGAGCGCGAGGAGGTGCTGCTGCTGCCCGG	1257
QY	1096	CGGACGACCTTACGAGAGATCTGCTGGACACCGCGGACTTACCGACATCGTGTGCGAG	1155
DB	1258	CGGACGACCTTACGAGAGATCTGCTGGACACCGCGGACTTACCGACATCGTGTGCGAG	1317
QY	1156	GTGACGACATCCGCGACGCGCATTTGCCATCGACTACGACCGCTAGAGGGCTATGCTTAC	1215
DB	1318	GTGACGACATCCGCGACGCGCATTTGCCATCGACTACGACCGCTAGAGGGCTATGCTTAC	1377
QY	1216	TGGACAGATGACGAGGTGCGGGCCATCCGAGGGCGTACTTGGACGGGTCTGGGGCGCGAG	1275
DB	1378	TGGACAGATGACGAGGTGCGGGCCATCCGAGGGCGTACTTGGACGGGTCTGGGGCGCGAG	1437
QY	1276	ACGCTGTGTAACACCGAGATCAACGACCCCGATGCGATTCGCGGTGCGATGCGGTGCGCGA	1335
DB	1438	ACGCTGTGTAACACCGAGATCAACGACCCCGATGCGATTCGCGGTGCGATGCGGTGCGCGA	1497
QY	1336	AACCTCTACTGGACCGGACGCGGACCGGACCGGATCGAGGTGACGCGCTCAACGGGACCC	1395
DB	1498	AACCTCTACTGGACCGGACGCGGACCGGACCGGATCGAGGTGACGCGCTCAACGGGACCC	1557
QY	1396	TCCCGCAAGATCTCTGTGTGCGAGGACCTTGAACGAGCCCGGAGCCATCGCACTGACCCCC	1455
DB	1558	TCCCGCAAGATCTCTGTGTGCGAGGACCTTGAACGAGCCCGGAGCCATCGCACTGACCCCC	1617
QY	1456	GTGATGGGCTCATGTACTGACAGACTGGGAGAGAAACCTTAAATCGAGTGTGCCAAC	1515
DB	1618	GTGATGGGCTCATGTACTGACAGACTGGGAGAGAAACCTTAAATCGAGTGTGCCAAC	1677
QY	1516	TTGATGGGCGAGGAGCGGCTGTGCTCAATGCTCTCCCTCGGGTGGCCCAACGGGCTG	1575
DB	1678	TTGATGGGCGAGGAGCGGCTGTGCTCAATGCTCTCCCTCGGGTGGCCCAACGGGCTG	1737
QY	1576	GCCTTGACCTGACAGGAGGAGAGTCTTATGCGGAGACCGCAAGACAGACAAATCGAG	1635
DB	1738	GCCTTGACCTGACAGGAGGAGAGTCTTATGCGGAGACCGCAAGACAGACAAATCGAG	1797
QY	1636	GTGATCAATCTGTGAGAGGAGCGGACCTCTCTGAGGAGCAAGCTTCCCGCAATT	1695
DB	1798	GTGATCAATCTGTGAGAGGAGCGGACCTCTCTGAGGAGCAAGCTTCCCGCAATT	1857
QY	1696	TTGGGTTTACGCTGCTGGGGGAGCTTCTATCTACTGAGCTGACTGGCAGCGCGGAGCATC	1755

Db 1858 TTCCGGTTTCA CGCTGCTGGGGACTTTCATCTACTGACTGACTGCGAGCGCGGAGCATC 1917
QY 1756 GAGCGGGTGCA CAAGGTTCAAGGCCAGCGCGAGCGTGCATCATGTGACAGCTGCCGACCTG 1815
Db 1918 GAGCGGGTGCA CAAGGTTCAAGGCCAGCGCGAGCGTGCATCATGTGACAGCTGCCGACCTG 1977
QY 1816 ATGGGGCTCA AAGCTGTGAATGTGGCCAAAGTGTGTCGGAACCAACCCGCTGTGGGACAGG 1875
Db 1978 ATGGGGCTCA AAGCTGTGAATGTGGCCAAAGTGTGTCGGAACCAACCCGCTGTGGGACAGG 2037
QY 1876 AACGGGGGTGAGCACCCTGTGCTTTTCACACCCACGAAACCCGCTGTGGCTGCCCTC 1935
Db 2038 AACGGGGGTG -AGCCACCTGTGCTTCTTCACACCCACGAAACCCGCTGTGGCTGCCCTC 2096
QY 1936 ATCCGCTTGGAGCTGCTGAGTGACATGAGAGCCTGTCATCGCTGAGGCCCTTCTTGCTC 1995
Db 2097 ATCCGCTTGGAGCTGCTGAGTGACATGAGAGCCTGTCATCGCTGAGGCCCTTCTTGCTC 2156
QY 1996 TTCAACAGCAGAGCGCCATCCACAGGATCTCCCTCGAGAGCAATAACACAGAGCTGGCC 2055
Db 2157 TTCAACAGCAGAGCGCCATCCACAGGATCTCCCTCGAGAGCAATAACACAGAGCTGGCC 2216
QY 2056 ATCCGCTTCA CGGGCGTCAAGGAGGCTCAGCGCTCAGCGCTTGTGATGTTCACCAACCCAC 2115
Db 2217 ATCCGCTTCA CGGGCGTCAAGGAGGCTCAGCGCTCAGCGCTTGTGATGTTCACCAACCCAC 2276
QY 2116 ATCTACTGGA CAGAGCTCAGCTGAGACATCAGCGCGCTTCTGAAACGGGAGCTCG 2175
Db 2277 ATCTACTGGA CAGAGCTCAGCGCTTGAAGACCAATCAGCGCGCTTCTGAAACGGGAGCTCG 2336
QY 2176 GTGAGACAGCTGTGGAGTGTGGCTTGAATACCCGAGGCGATGCGCGTTGACTGGATG 2235
Db 2337 GTGAGACAGCTGTGGAGTGTGGCTTGAATACCCGAGGCGATGCGCGTTGACTGGATG 2396
QY 2236 GGCAGAAACCTCTACTGGCCGCA CACTGGGACCAAGATCGAAGTGGCGCGCTGGAC 2295
Db 2397 GGCAGAAACCTCTACTGGCCGCA CACTGGGACCAAGATCGAAGTGGCGCGCTGGAC 2456
QY 2296 GGCAGTTC CGGCAAGTCTCGTGTGGAGGACTTGGACACCCGAGGTGCTGGCCCTG 2355
Db 2457 GGCAGTTC CGGCAAGTCTCGTGTGGAGGACTTGGACACCCGAGGTGCTGGCCCTG 2516
QY 2356 GATCCCAACCAAGGCTACATCTACTGGACCGAGTGGGCGCAAGCCGAGGATCGTCGG 2415
Db 2517 GATCCCAACCAAGGCTACATCTGACCCGAGTGGGCGCAAGCCGAGGATCGTCGG 2576
QY 2416 GCCTTTCATGGA CGGGAACCACTGATGACGCTGTGGACAAAGGTGGGCGGCGCAACGAC 2475
Db 2577 GCCTTTCATGGA CGGGAACCACTGATGACGCTGTGGACAAAGGTGGGCGGCGCAACGAC 2636
QY 2476 CTCACCATGACTAGCTGACCGAGCTCTACTGGACCGAGCTTGGACACCAACATGATC 2535
Db 2637 CTCACCATGACTAGCTGACCGAGCTCTACTGGACCGAGCTTGGACACCAACATGATC 2696
QY 2536 GAGTGTCTCAACATGTGGGTGAGAGCGGTGTGATTGGCGAGCTCTCCGACCCG 2595
Db 2697 GAGTGTCTCAACATGTGGGTGAGAGCGGTGTGATTGGCGAGCTCTCCGACCCG 2756
QY 2596 TTCGGTCTGACGAGTACAGCGATTATATCTACTGGA CAGACTGGAATCTGCACAGCAT 2655
Db 2757 TTCGGTCTGACGAGTACAGCGATTATATCTACTGGA CAGACTGGAATCTGCACAGCAT 2816
QY 2656 GAGCGGGCGCA CAGACTAGCGGCGGACCGACCTCTATCCAGGGCCACCTGGACTTC 2715
Db 2817 GAGCGGGCGCA CAGACTAGCGGCGGACCGACCTCTATCCAGGGCCACCTGGACTTC 2876
QY 2716 GTGATGACATCTCTGGTGTTCACCTCTCCGCGCAGGATGCGCTCAATGACTGTATGCAC 2775
Db 2877 GTGATGACATCTCTGGTGTTCACCTCTCCGCGCAGGATGCGCTCAATGACTGTATGCAC 2936
QY 2776 AACCAACGGGAGTGTGGGACGCTGTGCTTGCATCCCGCGGCGCACCGCTGGCGCTGC 2835

Db 2937 AACCAACGGGAGTGTGGGACGCTGTGCTTGCATCCCGCGGCGCACCGCTGGCGCTGC 2996
QY 2836 GCCTTCA CACTACACCTTGAGACCCAGAGCCGGAACCTGACAGCCCGCCACCACTTCTTG 2895
Db 2997 GCCTTCA CACTACACCTTGAGACCCAGAGCCGGAACCTGACAGCCCGCCACCACTTCTTG 3056
QY 2896 CTGTTTCAGCCAGAAATCTGCCATCAGTCCGATGATCCCGACGACCAAGCAGCAGCCGAT 2955
Db 3057 CTGTTTCAGCCAGAAATCTGCCATCAGTCCGATGATCCCGACGACCAAGCAGCAGCCGAT 3116
QY 2956 CTGATCTGCCCCCTGATGACTGAGAAAGTCAAGACCTGACTATGACCCACTGAC 3015
Db 3117 CTGATCTGCCCCCTGATGACTGAGAAAGTCAAGACCTGACTATGACCCACTGAC 3176
QY 3016 AAGTTTCATCTACTGGGTGATGGGGCCAGAAACATCAAGCAGCAGCAAGGACGAGGAC 3075
Db 3177 AAGTTTCATCTACTGGGTGATGGGGCCAGAAACATCAAGCAGCAGCAAGGACGAGGAC 3236
QY 3076 CAGCCCTTTGTTTTGACCTCTCTGAGCCAAAGCCAAACCCAGACAGGACGCCCCACGAC 3135
Db 3237 CAGCCCTTTGTTTTGACCTCTCTGAGCCAAAGCCAAACCCAGACAGGACGCCCCACGAC 3296
QY 3136 CTGAGCATTCGACATCTA CAGCCGACACTGTTCTTGGACGTGCGAGGCAACAATACCATC 3195
Db 3297 CTGAGCATTCGACATCTA CAGCCGACACTGTTCTTGGACGTGCGAGGCAACAATACCATC 3356
QY 3196 AACGTTCACAGGCTGAGGGGAGCCATGGGGGTGCTGCTGGGGACGCGGACAG 3255
Db 3357 AACGTTCACAGGCTGAGGGGAGCCATGGGGGTGCTGCTGGGGACGCGGACAG 3416
QY 3256 CCCAGGCGCATCTGCTGCTCAACCGGAGCGAGGGTACCTGTACTTCA CCAACATGACGAC 3315
Db 3417 CCCAGGCGCATCTGCTGCTCAACCGGAGCGAGGGTACCTGTACTTCA CCAACATGACGAC 3476
QY 3316 CGGCAACCAAGATTCGAAACCGCAGCCCTGGA CCGCACCGGACGCGAGGTCTCTTTCAC 3375
Db 3477 CGGCAACCAAGATTCGAAACCGCAGCCCTGGA CCGCACCGGACGCGAGGTCTCTTTCAC 3536
QY 3376 ACCGGCTCATCCGCTGCTGGCCCTGCTGGTGA CACACACTGGGCAAGCTGTCTCG 3435
Db 3537 ACCGGCTCATCCGCTGCTGGCCCTGCTGGTGA CACACACTGGGCAAGCTGTCTCG 3596
QY 3436 GTGACCGCGACCTTGAACCGCATTTGAGAGCTGTGACCTGT CAGGGGCAACCGCTGAC 3495
Db 3597 GTGACCGCGACCTTGAACCGCATTTGAGAGCTGTGACCTGT CAGGGGCAACCGCTGAC 3656
QY 3496 CTGAGGACGCGCAACATCTGTCAGCCCTTGGGCTGACCAATCTTGGCAAGCATCTTAC 3555
Db 3657 CTGAGGACGCGCAACATCTGTCAGCCCTTGGGCTGACCAATCTTGGCAAGCATCTTAC 3716
QY 3556 TGGATCGACCGCCAGCAGCAGATGATCGAGGTGTGGAGAGACCA CCGGGGCAAGCGG 3615
Db 3717 TGGATCGACCGCCAGCAGCAGATGATCGAGGTGTGGAGAGACCA CCGGGGCAAGCGG 3776
QY 3616 ACTCGCATCCAGGGCCGTGTGCGCCACTCACTGGCATCCATG CAGTGGAGAGTCAAG 3675
Db 3777 ACTCGCATCCAGGGCCGTGTGCGCCACTCACTGGCATCCATG CAGTGGAGAGTCAAG 3836
QY 3676 CTGAGGAGTTCAGCCCAACCCATGTCGCGTGA CAAATGGTGGCTGCTCCCACTGT 3735
Db 3837 CTGAGGAGTTCAGCCCAACCCATGTCGCGTGA CAAATGGTGGCTGCTCCCACTGT 3896
QY 3736 ATTGCCAAGGTTGATGGGACACCAAGGTGTCA TGGCCAGTCCACCTGCTCTCGAG 3795
Db 3897 ATTGCCAAGGTTGATGGGACACCAAGGTGTCA TGGCCAGTCCACCTGCTCTCGAG 3956
QY 3796 AACCTGCTGACTGTGTGAGAGCGCCCACTGCTCCCGACCA CAGTTTGCATGTGCCACA 3855
Db 3957 AACCTGCTGACTGTGTGAGAGCGCCCACTGCTCCCGACCA CAGTTTGCATGTGCCACA 4016
QY 3856 GGGGAGATCGACTGTATCCCGGGGCTGCGCTGTGA CCGGCTTCCGAGTGGAGTAC 3915
Db 4017 GGGGAGATCGACTGTATCCCGGGGCTGCGCTGTGA CCGGCTTCCGAGTGGAGTAC 4076

QY 3916 CAGAGGACGAGGAGGCTGCCCGTGTCTCCGCGCCGCCAGTTCCTCCCTGCGCGCGGGGT 3975
Db 4077 CAGAGGACGAGGAGGAGGCTGCCCGGTG-GCTCCGCGCCGCCAGTTCCTCCCTGCGCGCGGGGT 4135
QY 3976 CAGTGTGTGAGCTCGCCCTGCGCTGCGAGCGGAGGAGAGTGTGAGGACCGCTCAGAC 4035
Db 4136 CAGTGTGTGAGCTCGCCCTGCGCTGCGAGCGGAGGAGAGTGTGAGGACCGCTCAGAC 4195
QY 4036 GAGGCGGACTGTGACGCCATCTGCTGCTCCCAACCAAGTTCGCGTGTGCGAGCGCCAGTGT 4095
Db 4196 GAGGCGGACTGTGACGCCATCTGCTGCTCCCAACCAAGTTCGCGTGTGCGAGCGCCAGTGT 4255
QY 4096 GTCCCTCATCAACAGAGTGTGAGTCTCTCCCGGAGTGTATGAGCGGTCCGACGAGTTC 4155
Db 4256 GTCCCTCATCAACAGAGTGTGAGTCTCTCCCGGAGTGTATGAGCGGTCCGACGAGTTC 4315
QY 4156 ATGTGTGAAATCACCAGCGCCCTCAGACGACAGCGCGGCCACAGAGTGTGCCATCGG 4215
Db 4316 ATGTGTGAAATCACCAGCGCCCTCAGACGACAGCGCGGCCACAGAGTGTGCCATCGG 4375
QY 4216 CCGGTCAATGGCATCATCTCTCTCTCTCTGTCATGAGTGTGTCTATTTTGTGTCCAG 4275
Db 4376 CCGGTCAATGGCATCATCTCTCTCTCTCTGTCATGAGTGTGTCTATTTTGTGTCCAG 4435
QY 4276 CCGGTGTGTGCGAGCGTATGCGGGGCGCAACGGGCGCTTCCCGCACAGTATGTACG 4335
Db 4436 CCGGTGTGTGCGAGCGTATGCGGGGCGCAACGGGCGCTTCCCGCACAGTATGTACG 4495
QY 4336 GGGACCCCGCAGTGTCCCTCAATTTTCATAGCCCCGGGGTTCACAGTGTGCGCCCTTC 4395
Db 4496 GGGACCCCGCAGTGTCCCTCAATTTTCATAGCCCCGGGGTTCACAGTGTGCGCCCTTC 4555
QY 4396 ACAGGCTGTGATGCGGAAAGTTCATGATGAGTCCGTGAGCTGTGATGCGGGGCGCGGGC 4455
Db 4556 ACAGGCTGTGATGCGGAAAGTTCATGATGAGTCCGTGAGCTGTGATGCGGGGCGCGGGC 4615
QY 4456 GGGGTGCGCTGTACGACCGGAGACAGCTCAGGGGCGCTGTGCGAGGAGTGTGCGAGC 4515
Db 4616 GGGGTGCGCTGTACGACCGGAGACAGCTCAGGGGCGCTGTGCGAGGAGTGTGCGAGC 4675
QY 4516 ACAGGCGCAGCTGTACCGCGCATCTCTGAAACCGCGCGCTTCCCGGCCACGAGCCCC 4575
Db 4676 ACAGGCGCAGCTGTACCGCGCATCTCTGAAACCGCGCGCTTCCCGGCCACGAGCCCC 4735
QY 4576 TCCCTGTACACATGGAATGTTTCTCTTCAACATTCGCGGCATTCGCGAGACCGTAC 4635
Db 4736 TCCCTGTACACATGGAATGTTTCTCTTCAACATTCGCGGCATTCGCGAGACCGTAC 4795
QY 4636 AGGCCCTACATCTCGAGGAATGCGCGCCCGGAGAGTACTTCCATCTCTTCCCGCCCTCCG 4695
Db 4796 AGGCCCTACATCTCGAGGAATGCGCGCCCGGAGAGTACTTCCATCTCTTCCCGCCCTCCG 4855
QY 4696 GACAGCGACTACAGCGCGCGCTGGAAGCGCAGTACTACTCTGATTTGAATTCG 4755
Db 4856 GACAGCGACTACAGCGCGCGCTGGAAGCGCAGTACTACTCTGATTTGAATTCG 4915
QY 4756 GACTCAGACCCCTATCAGACCCGACCCAGCGCCACAGCGAGTACTGTGCGGCGAGAC 4815
Db 4916 GACTCAGACCCCTATCAGACCCGACCCAGCGCCACAGCGAGTACTGTGCGGCGAGAC 4975
QY 4816 AGCTGCCCGCGCTGCGCGCCACCGAGAGGAGTACTTCCATCTCTTCCCGCCCTCCG 4875
Db 4976 AGCTGCCCGCGCTGCGCGCCACCGAGAGGAGTACTTCCATCTCTTCCCGCCCTCCG 5035
QY 4876 TCCCGCTGACAGGACTATCTGACTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 4935
Db 5036 TCCCGCTGACAGGACTATCTGACTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 5095
QY 4936 TAAATAGTTTAAATATGAACAAAGAAATATATTTTATGATTTTAAATATATATAT 4995
Db 5096 TAAATAGTTTAAATATGAACAAAGAAATATATTTTATGATTTTAAATATATATAT 5155

QY 4996 AATTGGGATTTTAAAAACATGAGAAATGTGAATGTGATGGGTGGGCGAGGCTGGGAGA 5055
Db 5156 AATTGGGATTTTAAAAACATGAGAAATGTGAATGTGATGGGTGGGCGAGGCTGGGAGA 5215
QY 5056 ACTTTGTA 5063
Db 5216 ACTTTGTA 5223
RESULT 12
US-09-060-299-26
; Sequence 26, Application US/09060299
; Patent No. 6545137
; GENERAL INFORMATION:
; APPLICANT: Todd, John A
; APPLICANT: Hess, John W
; APPLICANT: Caskey, Charles T
; APPLICANT: Cox, Roger D
; APPLICANT: Gerhold, David
; APPLICANT: Hammond, Holly
; APPLICANT: Hey, Patricia
; APPLICANT: Kawaguchi, Yoshihiko
; APPLICANT: Merriman, Tony R
; APPLICANT: Metzker, Michael L
; TITLE OF INVENTION: No. 6545137el Receptor
; NUMBER OF SEQUENCES: 455
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon and Vanderhye
; STREET: 1100 No. 6545137th Glebe Road, Eighth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: US
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/060,299
; FILING DATE: 15-APR-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/043,553
; FILING DATE: 15-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/048,740
; FILING DATE: 05-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: B.J. Sadoff
; REGISTRATION NUMBER: 36,663
; REFERENCE/DOCKET NUMBER: 620-35
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)816-4091
; TELEFAX: (703)816-4100
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5125 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-060-299-26
Query Match 36.6%; Score 4890.6; DB 4; Length 5125;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 4915; Conservative 0; Mismatches 4; Indels 2; Gaps 2;
QY 143 CCTCGCGCTCTCTGCTATTTTCCAAACCGCGGAGACGTACGGCTGGTGGAGCGCGCGGAG 202
Db 167 CCTCGCGCTCTCTGCTATTTTCCAAACCGCGGAGACGTACGGCTGGTGGAGCGCGCGGAG 226
QY 203 TCAAGCTGGAGTCCACCATCGTGTACGCGCTGGAGGATGCGGCGCAGTGTGACTTCC 262
Db 227 TCAAGCTGGAGTCCACCATCGTGTACGCGCTGGAGGATGCGGCGCAGTGTGACTTCC 286

263 AGTTTTCCAGGAGCGGTGTAATGGAACAGAGCTGAGCGAGGAGCCATCAAGCAGACCT 322
| | | | |
287 AGTTTTCCAGGAGCGGTGTAATGGAACAGAGCTGAGCGAGGAGCCATCAAGCAGACCT 346
| | | | |
323 ACCTGAACAGACGCGGCGCCGCTGCAAGACGTGGTCAATCTCCGGCCCTGCTCTCCCG 382
| | | | |
347 ACCTGAACAGACGCGGCGCCGCTGCAAGACGTGGTCAATCTCCGGCCCTGCTCTCCCG 406
| | | | |
383 ACGGCCTCGCTGCACTGGGTGGGCAAGAGCTGTACTGACCGGACTCAGAGACCAACC 442
| | | | |
407 ACGGCCTCGCTGCACTGGGTGGGCAAGAGCTGTACTGACCGGACTCAGAGACCAACC 466
| | | | |
443 GCATCGAGGTGGCCAACTCAATGGCAATCCCGAAGGTGCTCTTCTGGCAGACCTTG 502
| | | | |
467 GCATCGAGGTGGCCAACTCAATGGCAATCCCGAAGGTGCTCTTCTGGCAGACCTTG 526
| | | | |
503 ACCAGCCTTAGGCGCCATCGCTTGACCCCGCTCAACGGGTACATGTACTGGACAGCTGGG 562
| | | | |
527 ACCAGCGAGGCGCCATCGCTTGACCCCGCTCAACGGGTACATGTACTGGACAGCTGGG 586
| | | | |
563 GTGAGACGCCCCGGATTGAGCGGCGAGGGATGGATGGCAGCACCGGAGAGATCATTTGG 622
| | | | |
587 GTGAGACGCCCCGGATTGAGCGGCGAGGGATGGATGGCAGCACCGGAGAGATCATTTGG 646
| | | | |
623 ACTCGGACATTTACTTGGCCCAATGGACTGACCATCGACTGGAGGAGCAGAGGCTCTACT 682
| | | | |
647 ACTCGGACATTTACTTGGCCCAATGGACTGACCATCGACTGGAGGAGCAGAGGCTCTACT 706
| | | | |
683 GGGCTGACGCGCAAGCTCAGCTTTCAATCCACCGTGCACACTGGACGGGTCTGTTCCGGCAGA 742
| | | | |
707 GGGCTGACGCGCAAGCTCAGCTTTCAATCCACCGTGCACACTGGACGGGTCTGTTCCGGCAGA 766
| | | | |
743 AGTGTGTGGAGGCGAGCTGACGACACCGCTTTCGCGCTGACGCTCTCCGGGAGACCTCTGT 802
| | | | |
767 AGTGTGTGGAGGCGAGCTGACGACACCGCTTTCGCGCTGACGCTCTCCGGGAGACCTCTGT 826
| | | | |
803 ACTGACAGACTGCGAGACCGCTTTCATCCATGCTGCAACAGCGCACTGGGGGGGAGA 862
| | | | |
827 ACTGACAGACTGCGAGACCGCTTTCATCCATGCTGCAACAGCGCACTGGGGGGGAGA 886
| | | | |
863 GGAAGGAGATCCTGAGTGCCCTTACTCACCCATGGACATCCAGGTGCTGAGCCAGGAGC 922
| | | | |
887 GGAAGGAGATCCTGAGTGCCCTTACTCACCCATGGACATCCAGGTGCTGAGCCAGGAGC 946
| | | | |
923 GGCAGCCTTTCTTCCACACTCGCTGTGAGGAGGACAAATGGCGGCTGTCTCCACCTGTGCC 982
| | | | |
947 GGCAGCCTTTCTTCCACACTCGCTGTGAGGAGGACAAATGGCGGCTGTCTCCACCTGTGCC 1006
| | | | |
983 TGCTGTCCCCAAGCGAGCCTTTCTACACATCGCGCTGCCCCACCGGTGTGAGCTGCAGG 1042
| | | | |
1007 TGCTGTCCCCAAGCGAGCCTTTCTACATCGCGCTGCCCCACCGGTGTGAGCTGCAGG 1066
| | | | |
1043 ACAACGCGCAGGACGTGTAAGCAGGAGCCGAGGAGGTGCTGCTGCTGGCCCGCGGACGG 1102
| | | | |
1067 ACAACGCGCAGGACGTGTAAGCAGGAGCCGAGGAGGTGCTGCTGCTGGCCCGCGGACGG 1126
| | | | |
1103 ACCTACGAGGATCTCGCTGAGACAGCGCGGACTTCAACGACATCGTGTGCTGAGGTGAGCG 1162
| | | | |
1127 ACCTACGAGGATCTCGCTGAGACAGCGCGGACTTTCACGACATCGTGTGCTGAGGTGAGCG 1186
| | | | |
1163 ACATCCGCGCAGCCATTGGCAATGCACTACGACCCGCTAGAGGGGTATGTCTACTTGGACAG 1222
| | | | |
1187 ACATCCGCGCAGCCATTGGCAATGCACTACGACCCGCTAGAGGGGTATGTCTACTTGGACAG 1246
| | | | |
1223 ATGACGAGGTGCGGCGCCATCCGACGGGCGTACTTGGACGGGTCTGGGCGCAGACGCTGG 1282
| | | | |
1247 ATGACGAGGTGCGGCGCCATCCGACGGGCGTACTTGGACGGGTCTGGGCGCAGACGCTGG 1306
| | | | |
1283 TCAACACGAGATCAACGACCCCGATGGCATCGCGGTGCACTGGGTGGCCCGGAAACCTCT 1342
| | | | |
1307 TCAACACGAGATCAACGACCCCGATGGCATCGCGGTGCACTGGGTGGCCCGGAAACCTCT 1366
| | | | |

1343 ACTGACCGACACGCGGACCGGACCGCATCGAGGTGACGCGCTCAACGCGACCTCCCGCA 1402
| | | | |
1367 ACTGACCGACACGCGGACCGGACCGCATCGAGGTGACGCGCTCAACGCGACCTCCCGCA 1426
| | | | |
1403 AGATCTCTGGTGTGCGAGGACCTGACGAGCGCCCGAGCCATCGCACTGCAACCCCGTGATGG 1462
| | | | |
1427 AGATCTCTGGTGTGCGAGGACCTGACGAGCGCCCGAGCCATCGCACTGCAACCCCGTGATGG 1486
| | | | |
1463 GCCTCATGTACTGACAGACTGGGGAGAGACCTTAAATCGAGTGTGCCAATTTGGATG 1522
| | | | |
1487 GCCTCATGTACTGACAGACTGGGGAGAGAACCCCTAAATCGAGTGTGCCAATTTGGATG 1546
| | | | |
1523 GGCAGGAGCGCGTGTGCTGGTCAATGCTCCCTCGGTGGCCCAACGCGCTGGCCCTGG 1582
| | | | |
1547 GGCAGGAGCGCGTGTGCTGGTCAATGCTCCCTCGGTGGCCCAACGCGCTGGCCCTGG 1606
| | | | |
1583 ACCTGCAAGGAGGGGAAGCTTACTGGGGAGACGCAAGACAGACAAGATCGAGGTGATCA 1642
| | | | |
1607 ACCTGCAAGGAGGGGAAGCTTACTGGGGAGACGCAAGACAGACAAGATCGAGGTGATCA 1666
| | | | |
1643 ATGTTGATGGAGCGAAGAGGCGGACCTCTCTGGAGGACAGCTCCGCGACATTTTCGGGT 1702
| | | | |
1667 ATGTTGATGGAGCGAAGAGGCGGACCTCTCTGGAGGACAGCTCCGCGACATTTTCGGGT 1726
| | | | |
1703 TCACGCTGCTGGGGGACTTCACTACTTGGACTGACTGGCAGCGCGCAGCATCGAGCGGG 1762
| | | | |
1727 TCACGCTGCTGGGGGACTTCACTACTTGGACTGACTGGCAGCGCGCAGCATCGAGCGGG 1786
| | | | |
1763 TGCAACAAGGTCAAGGCGAGCGGACGTCATCATTTGACCAAGCTCCCGACCTGATGGGGC 1822
| | | | |
1787 TGCAACAAGGTCAAGGCGAGCGGACGTCATCATTTGACCAAGCTCCCGACCTGATGGGGC 1846
| | | | |
1823 TCACAAGGTGGAATGTGGCCCAAGSTCTCGGAACTCAACCCGCTGTGCGGACAGAACGGG 1882
| | | | |
1847 TCACAAGGTGGAATGTGGCCCAAGSTCTCGGAACTCAACCCGCTGTGCGGACAGAACGGG 1906
| | | | |
1883 GGTGACGACCTGTGTCTTTCACACCCCAACGCAACCGGTGTGGCTGCCCACTCGGCCC 1942
| | | | |
1907 GGTGACGACCTGTGTCTTTCACACCCCAACGCAACCGGTGTGGCTGCCCACTCGGCCC 1966
| | | | |
1943 TGGAGCTGTGAGTGAACATGAAGACCTGCAATGTCCTGAGGCGCTTCTTGTGCTTTCACCA 2002
| | | | |
1967 TGGAGCTGTGAGTGAACATGAAGACCTGCAATGTCCTGAGGCGCTTCTTGTGCTTTCACCA 2026
| | | | |
2003 GCAGAGCGGCCNTCCACAGGATCTCCCTCGAGACCAATTAACAAGCTGGCCCATCCCGC 2062
| | | | |
2027 GCAGAGCGGCCATCCACAGGAT - TCCCTCGAGACCAATTAACAAGCTGGCCCATCCCGC 2085
| | | | |
2063 TCAGGCGCTCAAGGAGGCGCTCAGCCCTGGAATTTGATGTGTCCAAACAACACATCTACT 2122
| | | | |
2086 TCAGGCGCTCAAGGAGGCGCTCAGCCCTGGAATTTGATGTGTCCAAACAACACATCTACT 2145
| | | | |
2123 GGCAGACGTCAGCCTGAAGACCATCAGCGCGCTTCATGAACGGGAGCTCGGTGGAGC 2182
| | | | |
2146 GGCAGACGTCAGCCTGAAGACCATCAGCGCGCTTCATGAACGGGAGCTCGGTGGAGC 2205
| | | | |
2183 ACCTGTGTGAGTGTGGCCTTGAATACCCCGAGGCGATGGCGGTGCACTGGATGGGACAGA 2242
| | | | |
2206 ACCTGTGTGAGTGTGGCCTTGAATACCCCGAGGCGATGGCGGTGCACTGGATGGGACAGA 2265
| | | | |
2243 ACCTCTACTGGGCGGACACACTGGGACCAACAGAACTGGAAGTGGCGGCTGGACGGGACGT 2302
| | | | |
2266 ACCTCTACTGGGCGGACACACTGGGACCAACAGAACTGGAAGTGGCGGCTGGACGGGACGT 2325
| | | | |
2303 TCCGGCAAGTCTCTGCTGAGAGGACATTGGAACAACCGAGGTCTGCTGGCCCTTGGATCCCA 2362
| | | | |
2326 TCCGGCAAGTCTCTGCTGAGAGGACATTGGAACAACCGAGGTCTGCTGGCCCTTGGATCCCA 2385
| | | | |
2363 CCAGGGCTACATCTACTGACCGGAGTGGGCGGCAAGCCGAGGATCGTGGGCGGCTTCA 2422
| | | | |
2386 CCAGGGCTACATCTACTGACCGAGTGGGCGGCAAGCCGAGGATCGTGGGCGGCTTCA 2445
| | | | |
2423 TGGACGGGACCAACTGCAATGACGCTGTGGACAAAGGTGGGCGGCGGCAACGACCTCACCA 2482
| | | | |

Db 947 GGCAGCCTTCTTCACACTCGCTGTGAGGAGCAATGGCGGCTGCTCCCACTGTGCC 1006
Qy 983 TGCTGTCCCAAGCAGAGCCTTTCTACATGTCGCTGCCCAAGCGGTGTGACGTGCAAG 1042
Db 1007 TGCTGTCCCAAGCAGAGCCTTTCTACATGTCGCTGCCCAAGCGGTGTGACGTGCAAG 1066
Qy 1043 ACAACGGCAGGACGTGTAAAGCAGAGCGAGAGGTGTGCTGTGCTGCCCGCGGAGCGG 1102
Db 1067 ACNACGGCAGGACGTGTAAAGCAGAGCGAGAGGTGTGCTGTGCTGCCCGCGGAGCGG 1126
Qy 1103 ACCTACGAGGATCTCGCTGTGGACACGCCGGAATTCACCGACATCGTGTGACAGGTGGAGC 1162
Db 1127 ACCTACGAGGATCTCGCTGTGGACACGCCGGAATTCACCGACATCGTGTGACAGGTGGAGC 1186
Qy 1163 ACATCCGGCAGCCCATTTGCCATCGACTAGACCCCTAGAGGGCTATGTCTACTGGACAG 1222
Db 1187 ACATCCGGCAGCCCATTTGCCATCGACTAGACCCCTAGAGGGCTATGTCTACTGGACAG 1246
Qy 1223 ATGACAGGTGGCGGCCATTCGACGGGCTACCTGACCGGCTCTGGGCGCAGACGCTGG 1282
Db 1247 ATGACAGGTGGCGGCCATTCGACGGGCTACCTGACCGGCTCTGGGCGCAGACGCTGG 1306
Qy 1283 TCAACACCGAGATCAACGACCCCGATGCGATGCGGTGCACTGGGTGGCGCCGAAACCTCT 1342
Db 1307 TCAACACCGAGATCAACGACCCCGATGCGATGCGGTGCACTGGGTGGCGCCGAAACCTCT 1366
Qy 1343 ACTGACCGCAGCAGCGGACCGGACCCGATCGAGGTGACCGGCTCAACGGGACCTCCCGCA 1402
Db 1367 ACTGACCGCAGCAGCGGACCGGACCCGATCGAGGTGACCGGCTCAACGGGACCTCCCGCA 1426
Qy 1403 AGATCCTGGTTCGAGGACCTGGACGAGCCCGAGCCATCGCATGCAACCCGCTGATGG 1462
Db 1427 AGATCCTGGTTCGAGGACCTGGACGAGCCCGAGCCATCGCATGCAACCCGCTGATGG 1486
Qy 1463 GCCTCATGTACTGGACAGACTGGGAGAGAACCTTAAATCGAGTGTGCCAATTTGGATG 1522
Db 1487 GCCTCATGTACTGGACAGACTGGGAGAGAACCTTAAATCGAGTGTGCCAATTTGGATG 1546
Qy 1523 GGCAGGAGCGGCTGTGCTGCTCAATGCTCCTCGGTTGGCCCAACGGCTTGGCCCTGG 1582
Db 1547 GGCAGGAGCGGCTGTGCTGCTCAATGCTCCTCGGTTGGCCCAACGGCTTGGCCCTGG 1606
Qy 1583 ACCTGACGAGGGGAGCTTACTTGGGAGAGCGCCAGACAGACAAGATCGAGGTGATCA 1642
Db 1607 ACCTGACGAGGGGAGCTTACTTGGGAGAGCGCCAGACAGACAAGATCGAGGTGATCA 1666
Qy 1643 ATGTTGATGGGACGAGAGCGGACCTTCTGGAGGACAAGTCCCGCACATTTTCGGGT 1702
Db 1667 ATGTTGATGGGACGAGAGCGGACCTTCTGGAGGACAAGTCCCGCACATTTTCGGGT 1726
Qy 1703 TCACGCTGCTGGGAGCTTCACTACTGGAAGTGTGAGTGTGAGGCGCGGCGGATCGAGCGG 1762
Db 1727 TCACGCTGCTGGGAGCTTCACTACTGGAAGTGTGAGTGTGAGGCGCGGCGGATCGAGCGG 1786
Qy 1763 TGCAACAGGTCAAGGCCAGCGGAGCGCTCATCTTGAACAGCTGCGGACCTGTATGGGCG 1822
Db 1787 TGCAACAGGTCAAGGCCAGCGGAGCGCTCATCTTGAACAGCTGCGGACCTGTATGGGCG 1846
Qy 1823 TCAAAGCTGTGAATGTGCGCAAGGTGTGTCGGAACAACCCGTGTGCGGACAGGAACGGG 1882
Db 1847 TCAAAGCTGTGAATGTGCGCAAGGTGTGTCGGAACAACCCGTGTGCGGACAGGAACGGG 1906
Qy 1883 GGTGACGACCTGTGCTTCTTCAACCCCAAGCAACCGGTGTGGCTGCGCCCATCGGCC 1942
Db 1907 GGTGACGACCTGTGCTTCTTCAACCCCAAGCAACCGGTGTGGCTGCGCCCATCGGCC 1966
Qy 1943 TGGAGCTGCTGAGTGACATGAAGACCTGCATCTGCTGCTGAGGCTTCTTGTCTTCAACA 2002
Db 1967 TGGAGCTGCTGAGTGACATGAAGACCTGCATCTGCTGCTGAGGCTTCTTGTCTTCAACA 2026
Qy 2003 GCAGAGCGCCATCAAGGATCTCCCTCGAGACAACAATAACAAGACGTGGCCATTCGGC 2062

Db 2027 GCAGAGCCGCCATCCACAGGAT-TCCCTCGAGAGCAATAACAACGACGTGGCCATCCCGC 2085
Qy 2063 TCACGGGGCTCAGAGGCGCTCAGCCCTGGACTTTGATGTGTCTCAACCAACATCTACT 2122
Db 2086 TCACGGGGCTCAGAGGCGCTCAGCCCTGGACTTTGATGTGTCTCAACCAACATCTACT 2145
Qy 2123 GGACAGAGCTCAGCCTGAAGACCATCAGCCGCTTCATGAACGGGAGCTCGGTGGAGC 2182
Db 2146 GGACAGAGCTCAGCCTGAAGACCATCAGCCGCTTCATGAACGGGAGCTCGGTGGAGC 2205
Qy 2183 ACCTGTGAGGATTTGGCCTTGACTACCCCGAGGGCATGCGCTTGAAGTGGGCAAGA 2242
Db 2206 ACCTGTGAGGATTTGGCCTTGACTACCCCGAGGGCATGCGCTTGAAGTGGGCAAGA 2265
Qy 2243 ACCTTACTTGGGCGGACACTGGGACCAACAGAAATCGAAATGCGCGGCTGAGCGGCGAGT 2302
Db 2266 ACCTTACTTGGGCGGACACTGGGACCAACAGAAATCGAAATGCGCGGCTGAGCGGCGAGT 2325
Qy 2303 TCCGCAAGTCTCTGTGTGGAGGACTTTGGACAAACCCGAGGTTCGCTGGCCCTGGATCCCA 2362
Db 2326 TCCGCAAGTCTCTGTGTGGAGGACTTTGGACAAACCCGAGGTTCGCTGGCCCTGGATCCCA 2385
Qy 2363 CCAAGGGCTACATCTACTGGAACCGGCGGCAACCGAGGATCGTGGCGGCTTCA 2422
Db 2386 CCAAGGGCTACATCTACTGGAACCGGCGGCAACCGAGGATCGTGGCGGCTTCA 2445
Qy 2423 TGGACGGGACCAATGTCATGACGCTGTGGACAAAGTGGGCGGCGGCAACGACCTCAACA 2482
Db 2446 TGGACGGGACCAATGTCATGACGCTGTGGACAAAGTGGGCGGCGGCAACGACCTCAACA 2505
Qy 2483 TTGACTACGCTGACAGCGCTCTACTTGGACCGGACTTGGACCAACATGATCGAGTCTGT 2542
Db 2506 TTGACTACGCTGACAGCGCTCTACTTGGACCGGACTTGGACCAACATGATCGAGTCTGT 2565
Qy 2543 CCAACATGCTGGGTCAAGAGCGGCTGTGATTTGGCGAGATCTCCCGCACCCGTTCCGTC 2602
Db 2566 CCAACATGCTGGGTCAAGAGCGGCTGTGATTTGGCGAGATCTCCCGCACCCGTTCCGTC 2625
Qy 2603 TGGACGAGTACAGGATTTATATCTACTTGGACAGCTGGAATCTGCAAGCATTTGAGCGG 2662
Db 2626 TGGACGAGTACAGGATTTATATCTACTTGGACAGCTGGAATCTGCAAGCATTTGAGCGG 2685
Qy 2663 CCGACAGACTTAGCGGCGGCAACCGCACCTTATCCAGGCGCACCTGACCTTCTGTGATGG 2722
Db 2686 CCGACAGACTTAGCGGCGGCAACCGCACCTTATCCAGGCGCACCTGACCTTCTGTGATGG 2745
Qy 2723 ACATCTGGTGTTCATCTCTCCCGCAGAGATGCGCTCAATGATGTATGCAACAACG 2782
Db 2746 ACATCTGGTGTTCATCTCTCCCGCAGGATGCGCTCAATGATGTATGCAACAACG 2805
Qy 2783 GGCAGTGTGGCAGCTGTGCTTGGCATCCCGGCGGCGCACCGCTGGCGCTCCAC 2842
Db 2806 GGCAGTGTGGCAGCTGTGCTTGGCATCCCGGCGGCGCACCGCTGGCGCTCCAC 2865
Qy 2843 ACTACACCTTGGACCCAGCAGCGGCAACTGACGCGGCGGCGCACCCCTTCTGTGTTCA 2902
Db 2866 ACTACACCTTGGACCCAGCAGCGGCAACTGACGCGGCGGCGCACCCCTTCTGTGTTCA 2925
Qy 2903 GCCAGAAATCTGCCATCATGTGCGGATGATCCCGGACGACGACAGCCCGGATCTATCC 2962
Db 2926 GCCAGAAATCTGCCATCATGTGCGGATGATCCCGGACGACGACAGCCCGGATCTATCC 2985
Qy 2963 TGCCCTTGCATGGAAGTGGGAACTGAAAGCATGCAATGACCTGACCAAGTCA 3022
Db 2986 TGCCCTTGCATGGAAGTGGGAACTGAAAGCATGCAATGACCTGACCAAGTCA 3045
Qy 3023 TCTACTGGGTGGATGGGCGCGCAGAACATCAAGCGAGCGCAAGCGGAGCCCAAGCCCT 3082
Db 3046 TCTACTGGGTGGATGGGCGCGCAGAACATCAAGCGAGCGCAAGCGGAGCCCAAGCCCT 3105
Qy 3083 TTGTTTTGACCTCTCTGAGCGCAAGGCGCAAAACCCAGACAGGCGCCCAACGACCTCAGCA 3142
Db 3106 TTGTTTTGACCTCTCTGAGCGCAAGGCGCAAAACCCAGACAGGCGCCCAACGACCTCAGCA 3165

APPLICANT: Kawaguchi, Yoshihiko
APPLICANT: Merriman, Tony R
APPLICANT: Metzker, Michael L
TITLE OF INVENTION: No. 6545137el Receptor
NUMBER OF SEQUENCES: 455
CORRESPONDENCE ADDRESS:
ADDRESSES: Nixon and Vanderhye
STREET: 1100 No. 6545137th Glebe Road, Eighth Floor
City: Arlington
State: Virginia
COUNTRY: US
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/060,299
FILING DATE: 15-APR-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/043,553
FILING DATE: 15-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/048,740
FILING DATE: 05-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: B.J. Sadoff
TELEPHONE: (703) 816-4091
TELEFAX: (703) 816-4100
REGISTRATION NUMBER: 36,663
REFERENCE/DOCKET NUMBER: 620-35
TELECOMMUNICATION INFORMATION:
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 5166 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-060-299-23

Query Match 96.6%; Score 4890; DB 4; Length 5166;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 4915; Conservative 0; Mismatches 5; Indels 2; Gaps 2;

QY	142	GCCTCGCGCTCCTGCTATTTGCCAACCGCGGACGTACGGCTGGTGAGCGCGCGGA	201
DB	217	GCCTCGCGCTCCTGCTATTTGCCAACCGCGGACGTACGGCTGGTGAGCGCGCGGA	276
QY	202	GTCAAGCTGGAGTCCACCATCGTGTGTCAGCGGCTGGAGGATCGCGCGCAGTGGACTTC	261
DB	277	GTCAAGCTGGAGTCCACCATCGTGTGTCAGCGGCTGGAGGATCGCGCGCAGTGGACTTC	336
QY	262	CAGTTTTCAGAGGAGCGGTGTACTGGAACAGCGTGGAGGAGGCCATCAAGCAGACC	321
DB	337	CAGTTTTCAGAGGAGCGGTGTACTGGAACAGCGTGGAGGAGGCCATCAAGCAGACC	396
QY	322	TACCTGACACAGAGGGGGCGCGTGCAGAACGTGGTTCATCTCGGCTGGTCTCTCCC	381
DB	397	TACCTGACACAGAGGGGGCGCGGTGCAGAACGTGGTTCATCTCGGCTGGTCTCTCCC	456
QY	382	GACGCTCTCGCTCGGCTGGCTGGGTCAGAGGCTGTACTGGAAGGACTCAGAGACCAAC	441
DB	457	GACGCTCTCGCTCGGCTGGCTGGGTCAGAGGCTGTACTGGAAGGACTCAGAGACCAAC	516
QY	442	CGCATCGAGGTGGCCAACTCAATGGCAATCCCGGAAGGTGCTCTTCGCGAGGACCTT	501
DB	517	CGCATCGAGGTGGCCAACTCAATGGCAATCCCGGAAGGTGCTCTTCGCGAGGACCTT	576
QY	502	GACAGGCTAGGGCCATCGCTTGGACCCGCTCAGCGGTACATGTACTTGGACAGACTGG	561
DB	577	GACAGGCTAGGGCCATCGCTTGGACCCGCTCAGCGGTACATGTACTTGGACAGACTGG	636

QY	562	GGTGAGACGCGCGGATTGAGCGGGCAGGGATGGATGGCAGCACCCGGAAGATCATTTGTG	621
DB	637	GGTGAGACGCGCGGATTGAGCGGGCAGGGATGGATGGCAGCACCCGGAAGATCATTTGTG	696
QY	622	GACTCGGACATTTACTGCGCCCAATGGAGCTGACCATCGACCTGGAGGAGCAGAGCTCTAC	681
DB	697	GACTCGGACATTTACTGCGCCCAATGGAGCTGACCATCGACCTGGAGGAGCAGAGCTCTAC	756
QY	682	TGGGCTGACGCGCAAGCTCAGCTTCATCCACGCTGCCAACCTGGAGCGGCTCGTTCCGGCAG	741
DB	757	TGGGCTGACGCGCAAGCTCAGCTTCATCCACGCTGCCAACCTGGAGCGGCTCGTTCCGGCAG	816
QY	742	AAGTGTGGAGGGCAGCTGACGCAACCTTCGCTTCGCTGACGCTCTCCGGGAGCACTCTG	801
DB	817	AAGTGTGGAGGGCAGCTGACGCAACCTTCGCTTCGCTGACGCTCTCCGGGAGCACTCTG	876
QY	802	TACTGGACAGACTGGCAGACCGCTTCATCATGCTGCTCAACAGCGCACTGGGGGGAG	861
DB	877	TACTGGACAGACTGGCAGACCGCTTCATCATGCTGCTCAACAGCGCACTGGGGGGAG	936
QY	862	AGGAAGGAGATCCTGAGTGCCTCTACTCACCCATGGACATCCAGGTCTGAGCCAGAG	921
DB	937	AGGAAGGAGATCCTGAGTGCCTCTACTCACCCATGGACATCCAGGTCTGAGCCAGAG	996
QY	922	CGGAGACCTTTCTTCCACACTCGCTGTGAGGAGCAATGGGGCTGCTCCACCTGTGC	981
DB	997	CGGAGACCTTTCTTCCACACTCGCTGTGAGGAGCAATGGGGCTGCTCCACCTGTGC	1056
QY	982	CTGCTGTCCCCAAGCGAGCCCTTTCTACACATGCGCTTCCCGCCAGGGTGTGAGCTGCAG	1041
DB	1057	CTGCTGTCCCCAAGCGAGCCCTTTCTACACATGCGCTTCCCGCCAGGGTGTGAGCTGCAG	1116
QY	1042	GACAAAGGAGAGCTGTAAAGGAGGAGCGAGGAGTGTGCTGTGGCCCGGCGGAG	1101
DB	1117	GACAAAGGAGAGCTGTAAAGGAGGAGCGAGGAGTGTGCTGTGGCCCGGCGGAG	1176
QY	1102	GACCTAGCGAGGATCTCGCTGGACACGCGGACTTCAACGACATCTGCTCAGGTGAC	1161
DB	1177	GACCTAGCGAGGATCTCGCTGGACACGCGGACTTCAACGACATCTGCTCAGGTGAC	1236
QY	1162	GACATCCGGCAGCGCATTTGGCCATCGACTACGACCCGCTAGAGGCTATGTCTACTGACA	1221
DB	1237	GACATCCGGCAGCGCATTTGGCCATCGACTACGACCCGCTAGAGGCTATGTCTACTGACA	1296
QY	1222	GATGACGAGGTGCGGGCCATCCGACGCGCTACTGACGCGGTCTGGGCGCAGACGCTG	1281
DB	1297	GATGACGAGGTGCGGGCCATCCGACGCGCTACTGACGCGGTCTGGGCGCAGACGCTG	1356
QY	1282	GTCAACACCGAGATCAACGACCCCGATGGCATCGCGGTGCACTGGGTGGCCGGAACCTC	1341
DB	1357	GTCAACACCGAGATCAACGACCCCGATGGCATCGCGGTGCACTGGGTGGCCGGAACCTC	1416
QY	1342	TACTGGAACCGACACGCGCAGGACCGCATCGAGGTGACGCGCTTCAACGCGACCTCCCGC	1401
DB	1417	TACTGGAACCGACACGCGCAGGACCGCATCGAGGTGACGCGCTTCAACGCGACCTCCCGC	1476
QY	1402	AAGATCTGTGTGTCGGAGGACCTTGGACGCGCCGAGCCATCGCACTGCAACCCGCTGATG	1461
DB	1477	AAGATCTGTGTGTCGGAGGACCTTGGACGCGCCGAGCCATCGCACTGCAACCCGCTGATG	1536
QY	1462	GGCCTCATGTACTGGAACAGACTGGGGAGAGAACCTTAAATTCGAGTGTGCCAACTTGGAT	1521
DB	1537	GGCCTCATGTACTGGAACAGACTGGGGAGAGAACCTTAAATTCGAGTGTGCCAACTTGGAT	1596
QY	1522	GGGAGAGAGCGGTGTGTGCTCAATGCTTCCCTCGGGTGGCCCAAGCGGCTGGCCCTG	1581
DB	1597	GGGAGAGAGCGGTGTGTGCTCAATGCTTCCCTCGGGTGGCCCAAGCGGCTGGCCCTG	1656
QY	1582	GACCTGAGAGGGGAAAGCTTACTTGGGAGAGCGCCAAAGCAGACAAGATCAGGTGATC	1641
DB	1657	GACCTGAGAGGGGAAAGCTTACTTGGGAGAGCGCCAAAGCAGACAAGATCAGGTGATC	1716
QY	1642	AATGTTGATGGGACGAAAGAGCGGACCCCTCTCTGGAGGAGCAAGCTTCCGCGACATTTTCGGG	1701

Db 1717 AATGTTGATGGAGCAAGAGCGGACCTCTCTGGAGGACAAGCTCCCGCACATTTTCGGG 1776
QY TTACGCTGCTGGGGGACTTCACTACTACTGGAAGTGAAGTGGAGCGCGGAGCATCGAGCGG 1761
Db 1777 TTACGCTGCTGGGGGACTTCACTACTACTGGAAGTGAAGTGGAGCGCGGAGCATCGAGCGG 1836
QY GTGCAAAAGGTCAAGGCGAGCGGGAGCTCATATTGACCAAGCTGCGCGGAGCTGATGGGG 1821
Db 1837 GTGCACAAAGGTCAAGGCGAGCGGGAGCTCATATTGACCAAGCTGCGCGGAGCTGATGGGG 1896
QY 1822 CTCAAAGCTGGAATGTGGCCAAAGTCTGTGGAACCAACCGGTGTGGCGGACAGGAACGGG 1881
Db 1897 CTCAAAGCTGGAATGTGGCCAAAGTCTGTGGAACCAACCGGTGTGGCGGACAGGAACGGG 1956
QY 1882 GGTGTGACGCACTGTGTCTTTTCAACCCCAAGCAACCGGTGTGGCGGAGCTGCGCCATCGGC 1941
Db 1957 GGTGTGACGCACTGTGTCTTTTCAACCCCAAGCAACCGGTGTGGCGGAGCTGCGCCATCGGC 2016
QY 1942 CTGGAGCTGTGAGTGAATGAAGACCTGCACTGTGCTGAGGCGCTTCTTGGTCTTTCAAC 2001
Db 2017 CTGGAGCTGTGAGTGAATGAAGACCTGCACTGTGCTGAGGCGCTTCTTGGTCTTTCAAC 2075
QY 2002 AGCAGAGCGGCATCCACAGGATCTCCCTCGAGACCAATAACAAGAGCTGGCCATCCCG 2061
Db 2076 AGCAGAGCGGCATCCACAGGATCTCCCTCGAGACCAATAACAAGAGCTGGCCATCCCG 2135
QY 2062 CTCAGGGCGCTCAAGAGGCGCTCAGCCCTGAGCTTTGATGTGTCACAAACCAACCATCTAC 2121
Db 2136 CTCAGGGCGCTCAAGAGGCGCTCAGCCCTGAGCTTTGATGTGTCACAAACCAACCATCTAC 2195
QY 2122 TGGACAGAGCTCAGCTGGAAGACCATCAGCGCGCGCTTTCATGAACGGGAGCTCGGTGGAG 2181
Db 2196 TGGACAGAGCTCAGCTGGAAGACCATCAGCGCGCGCTTTCATGAACGGGAGCTCGGTGGAG 2255
QY 2182 CAGTGTGGAGTTTGGCTTGAATACCCCGAGGGCATGGCCGTTGATGATGGAATGGGCAAG 2241
Db 2256 CAGTGTGGAGTTTGGCTTGAATACCCCGAGGGCATGGCCGTTGATGATGGAATGGGCAAG 2315
QY 2242 AACCTTACTGGGCGGACACTGGGACCAACAGAAATCGAAGTGGCGGCTGGAGCGGCGAG 2301
Db 2316 AACCTTACTGGGCGGACACTGGGACCAACAGAAATCGAAGTGGCGGCTGGAGCGGCGAG 2375
QY 2302 TTCGGGCAAGTCTCGTGTGGAGGAGCTTGGACCAACCGAGGTGGCTGGGCGCTGGATCCC 2361
Db 2376 TTCGGGCAAGTCTCGTGTGGAGGAGCTTGGACCAACCGAGGTGGCTGGGCGCTGGATCCC 2435
QY 2362 ACCAAGGCTACATCTACTGACCCGAGTGGGCGGCAAGCCGAGATCGTGGGCGCTTC 2421
Db 2436 ACCAAGGCTACATCTACTGACCCGAGTGGGCGGCAAGCCGAGATCGTGGGCGCTTC 2495
QY 2422 ATGGACGGGACCAACTGACATGACCTGGTGGACAGGTGGGCGGCAAGCCGAGCTCAC 2481
Db 2496 ATGGACGGGACCAACTGACATGACCTGGTGGACAGGTGGGCGGCAAGCCGAGCTCAC 2555
QY 2482 ATTGACTACGCTGACCGGCTCTTACTGGACCGGACCTGGACACCAATGATCGAGTCG 2541
Db 2556 ATTGACTACGCTGACCGGCTCTTACTGGACCGGACCTGGACACCAATGATCGAGTCG 2615
QY 2542 TCCAAATGCTGGGTGAGGAGCGGTCGTGATGTCGACGATCTCCGCAACCGGTTCGGT 2601
Db 2616 TCCAAATGCTGGGTGAGGAGCGGTCGTGATGTCGACGATCTCCGCAACCGGTTCGGT 2675
QY 2602 CTGACGGAGTACAGGATATATCTACTGGACAGACTGGATCTGCACAGATTTGAGCGG 2661
Db 2676 CTGACGGAGTACAGGATATATCTACTGGACAGACTGGATCTGCACAGATTTGAGCGG 2735
QY 2662 GCCGACAGACTAGCGGCGGAAACCGCACCTCTATCCAGGGCCACCTGGACTTGGTATG 2721
Db 2736 GCCGACAGACTAGCGGCGGAAACCGCACCTCTATCCAGGGCCACCTGGACTTGGTATG 2795
QY 2722 GACATCTGTGGTTTCCACTCTCCCGCAGGATGGCTCAATGATGTATGCAACAAC 2781

Db 2796 GACATCTGTGGTTTCCACTCTCCCGCAGGATGGCTCAATGATGTATGACACAACAAC 2855
QY 2782 GGGCAGTGTGGGCGAGCTGTGCTTGGCCATCCCGCGCGCCACCGCTGGCGCTGCGCCTCA 2841
Db 2856 GGGCAGTGTGGGCGAGCTGTGCTTGGCCATCCCGCGCGCCACCGCTGGCGCTGCGCCTCA 2915
QY 2842 CACTACACCTTGGAGCCCGCAGCAGCGCAACTTGCAGCGCCGCCCAACCACTTCTTGTGTTTC 2901
Db 2916 CACTACACCTTGGAGCCCGCAGCAGCGCAACTTGCAGCGCCGCCCAACCACTTCTTGTGTTTC 2975
QY 2902 AGCCAGAAATCTGCATCAGTCCGATGATCCCGAGCAGCAGCAGCAGCAGCAGCAGCAGC 2961
Db 2976 AGCCAGAAATCTGCATCAGTCCGATGATCCCGAGCAGCAGCAGCAGCAGCAGCAGCAGC 3035
QY 2962 CTGCGCCCTGCACTGAGTGAAGAACTTCAAGCCATCGACTATGACCACTGAGCAAGTTTC 3021
Db 3036 CTGCGCCCTGCACTGAGTGAAGAACTTCAAGCCATCGACTATGACCACTGAGCAAGTTTC 3095
QY 3022 ATCTACTGGGTGGATGGGCGCAGAACTCAAGCGAGCAAGAGCAGCAGGACCCAGCCCC 3081
Db 3096 ATCTACTGGGTGGATGGGCGCAGAACTCAAGCGAGCAAGAGCAGCAGGACCCAGCCCC 3155
QY 3082 TTTGTTTGTGACCTCTCTGAGCCAAAGGCCAAACCCAGACAGCAGCAGCAGCAGCAGCAGC 3141
Db 3156 TTTGTTTGTGACCTCTCTGAGCCAAAGGCCAAACCCAGACAGCAGCAGCAGCAGCAGCAGC 3215
QY 3142 ATCGACATCTACACCGCGACACTGTTCTTGGACGTGCGAGGCGCAACCAATACCATCAACGTC 3201
Db 3216 ATCGACATCTACACCGCGACACTGTTCTTGGACGTGCGAGGCGCAACCAATACCATCAACGTC 3275
QY 3202 CACAGGCTGAGCGGGGAAGCCATGGGGTGGTGTGCTGGGGGACCGCGACAAAGCCAGG 3261
Db 3276 CACAGGCTGAGCGGGGAAGCCATGGGGTGGTGTGCTGGGGGACCGCGACAAAGCCAGG 3335
QY 3262 GCCATGCTGTCACACGGGAGGAGGTACTCTGTACTTTCACAAATGACAGACAGCGGGA 3321
Db 3336 GCCATGCTGTCACACGGGAGGAGGTACTCTGTACTTTCACAAATGACAGACAGCGGGA 3395
QY 3322 GCCAAGATCGAAACCGCGACCGCTCGACCGGACCGAGCGGAGTCTCTTTCACCAACCGCG 3381
Db 3396 GCCAAGATCGAAACCGCGACCGCTCGACCGGACCGAGCGGAGTCTCTTTCACCAACCGCG 3455
QY 3382 CTCATCCGCTGTGGCCCTGGTGGTGGACAAACAACAATGGGCAAGCTTCTTGGGTGGAC 3441
Db 3456 CTCATCCGCTGTGGCCCTGGTGGTGGACAAACAACAATGGGCAAGCTTCTTGGGTGGAC 3515
QY 3442 GCGGACCTGAAGCGCATTTGAGAGCTGTGACCTGTGAGGGGCAACCGCTGACCCCTGAG 3501
Db 3516 GCGGACCTGAAGCGCATTTGAGAGCTGTGACCTGTGAGGGGCAACCGCTGACCCCTGAG 3575
QY 3502 GAGCCCAACATCGTGCAGCTCTGGGCTGACCATCTTGGCAAGCATCTCTACTTGGATC 3561
Db 3576 GAGCCCAACATCGTGCAGCTCTGGGCTGACCATCTTGGCAAGCATCTCTACTTGGATC 3635
QY 3562 GACCGCAGCAGCAGATGATCGAGCTGTGGAGAGAACCAACCGGGGACAAGCGGACTCGC 3621
Db 3636 GACCGCAGCAGCAGATGATCGAGCTGTGGAGAGAACCAACCGGGGACAAGCGGACTCGC 3695
QY 3622 ATCCAGGGCGGTGTGGCCACCTCACTGGGCAATCCATGAGTGGAGGAGTCAAGCCTGAG 3681
Db 3696 ATCCAGGGCGGTGTGGCCACCTCACTGGGCAATCCATGAGTGGAGGAGTCAAGCCTGAG 3755
QY 3682 GAGTTCTCAGCCCAACATGTCGCGGTGACAAATGGTGGCTCTCCCATCTGTATTGCC 3741
Db 3756 GAGTTCTCAGCCCAACATGTCGCGGTGACAAATGGTGGCTCTCCCATCTGTATTGCC 3815
QY 3742 AAGGTTGATGGGACCAACACGGTGTCTATGCCAGTCCACCTCGTCTCTCGCAGAACCTG 3801
Db 3816 AAGGTTGATGGGACCAACACGGTGTCTATGCCAGTCCACCTCGTCTCTCGCAGAACCTG 3875
QY 3802 CTGACCTGTGGAGAGCGGCCCACTGTCTCCCGGACCAAGTTTGTGATGTGCCACAGGGGAG 3861
Db 3876 CTGACCTGTGGAGAGCGGCCCACTGTCTCCCGGACCAAGTTTGTGATGTGCCACAGGGGAG 3935

QY 3862 ATCGACTGATATCCCGGGGCGCTGGCGCTGTGACGGCTTTCCGAGTGGCATGACACAGAGC 3921
Db 3936 ATCGACTGATATCCCGGGGCGCTGGCGCTGTGACGGCTTTCCGAGTGGCATGACACAGAGC 3995
QY 3922 GACGAGGAGGCTGCCCGTGTGCTCCGCGCCAGTTCCCTGCGCGCGGGGTGAGTGT 3981
Db 3936 GACGAGGAGGCTGCCCGTGTGCTCCGCGCCAGTTCCCTGCGCGCGGGGTGAGTGT 4055
QY 3982 GTGGACCTGCGCTGCGTGGACGCGCGAGGAGAGACTGTGAGGACCGCTCAGACGAGCGG 4041
Db 4056 GTGGACCTGCGCTGCGTGGACGCGCGAGGAGAGACTGTGAGGACCGCTCAGAC-AGGCG 4114
QY 4042 GACTGTGACGCACTGCTGCTGCCGCAACAGTTCGGTGTGCGAGCGGCGAGTGTCTCTC 4101
Db 4115 GACTGTGACGCACTGCTGCTGCCGCAACAGTTCGGTGTGCGAGCGGCGAGTGTCTCTC 4174
QY 4102 ATCAACAGCAGTGGACTCTCTTCCCGACTGTATCGACGGCTCCGACGAGCTCATGTGT 4161
Db 4175 ATCAACAGCAGTGGACTCTCTTCCCGACTGTATCGACGGCTCCGACGAGCTCATGTGT 4234
QY 4162 GAAATCACCAAGCGCCCTCAGACGACAGCCCGGCCACAGCAGTGCATCGGGCCCGTC 4221
Db 4235 GAAATCACCAAGCGCCCTCAGACGACAGCCCGGCCACAGCAGTGCATCGGGCCCGTC 4294
QY 4222 ATTGGCATCATCTCTCTCTCTTGGTGTGAGTGTGATTTGTTGTGTGCGAGCGGTG 4281
Db 4295 ATTGGCATCATCTCTCTCTCTTGGTGTGAGTGTGATTTGTTGTGTGCGAGCGGTG 4354
QY 4282 GTGTGCCAGCGTATGCGGGGGCCAAAGCGGCCCTTCCCGCAGCAGTATGTACAGGGGACC 4341
Db 4355 GTGTGCCAGCGTATGCGGGGGCCAAAGCGGCCCTTCCCGCAGCAGTATGTACAGGGGACC 4414
QY 4342 CCGCAGTGTGCCCTCAATTTTATAGCCCCGGCGGTTCCTCAGCATGCGCCCTTCACAGGC 4401
Db 4415 CCGCAGTGTGCCCTCAATTTTATAGCCCCGGCGGTTCCTCAGCATGCGCCCTTCACAGGC 4474
QY 4402 ATCGCATCGGAAGTCCATGATGAGTCCGTGAGCCTGTATGGGGCGCGGGGGGTG 4461
Db 4475 ATCGCATCGGAAGTCCATGATGAGTCCGTGAGCCTGTATGGGGCGCGGGGGGTG 4534
QY 4462 CCCCTGTACGACCGGAACACAGTCCAGCGGGCTCGTCCAGCAGTCTGTCCAGACAGAG 4521
Db 4535 CCCCTGTACGACCGGAACACAGTCCAGCGGGCTCGTCCAGCAGTCTGTCCAGACAGAG 4594
QY 4522 GCCAGCTGTACCGCCGATCTGTAAACCCCGCGCTCCCGGCCACAGGACCCCTCCCTG 4581
Db 4595 GCCAGCTGTACCGCCGATCTGTAAACCCCGCGCTCCCGGCCACAGGACCCCTCCCTG 4654
QY 4582 TACAACATGGACATGTTCTACTCTTCAACATTCGGGCACTCGGAGACCGTACAGGCC 4641
Db 4655 TACAACATGGACATGTTCTACTCTTCAACATTCGGGCACTGTGAGACCGTACAGGCC 4714
QY 4642 TACATCATTCGAGGAATGGCGCCCGCCGACGAGCGCCCTGACGACCGAGTGTGACAGC 4701
Db 4715 TACATCATTCGAGGAATGGCGCCCGCCGACGAGCGCCCTGACGACCGAGTGTGACAGC 4774
QY 4702 GACTACAGCGCGCGCTGGAAGGCCAGCAAGTACTACTGGATTTGAACTCGGACTCA 4761
Db 4775 GACTACAGCGCGCGCTGGAAGGCCAGCAAGTACTACTGGATTTGAACTCGGACTCA 4834
QY 4762 GACCCCTATCCACCCGACCCGACCCGACCCGACCCGACCCGACCCGACCCGACCCGACCCG 4821
Db 4835 GACCCCTATCCACCCGACCCGACCCGACCCGACCCGACCCGACCCGACCCGACCCGACCCG 4894
QY 4822 CCGCCCTCGCCGACCCGACCCGACCCGACCCGACCCGACCCGACCCGACCCGACCCGACCCG 4881
Db 4895 CCGCCCTCGCCGACCCGACCCGACCCGACCCGACCCGACCCGACCCGACCCGACCCGACCCG 4954
QY 4882 TGCAGGAGTCACTGACCTTGGCGCGGCGCACTCTGCTGCTCTCTGTCGCCCTGTAATA 4941
Db 4955 TGCAGGAGTCACTGACCTTGGCGCGGCGCACTCTGCTGCTCTCTGTCGCCCTGTAATA 5014

QY 4942 GTTTTAAATATGAACAAAGAAAAATATATATTTATGATTTAAAAAATAAATATATTTGG 5001
Db 5015 GTTTTAAATATGAACAAAGAAAAATATATATTTATGATTTAAAAAATAAATATATTTGG 5074
QY 5002 GATTTTAAAAACATGAGAAATGTGAACCTGTGATGGGTGGCGAGGCTGGGAGAACTTTG 5061
Db 5075 GATTTTAAAAACATGAGAAATGTGAACCTGTGATGGGTGGCGAGGCTGGGAGAACTTTG 5134
QY 5062 TA 5063
Db 5135 TA 5136
RESULT 15
US-09-402-923A-23
; Sequence 23, Application US/09402923A
; Patent No. 6555654
; GENERAL INFORMATION:
; APPLICANT: Todd, John A
; Hess, John W
; Caskey, Charles T
; Cox, Roger D
; Gerhold, David
; Hammond, Holly
; Hey, Patricia
; Kawaguchi, Yoshihiko
; Merriman, Tony R
; Metzker, Michael L
; TITLE OF INVENTION: No. 6555654el LDL-Receptor
; NUMBER OF SEQUENCES: 455
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon and Vanderhye
; STREET: 1100 No. 6555654th Glebe Road, Eighth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: US
; ZIP: VA 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/402,923A
; FILING DATE: 14-Feb-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB98/01102
; FILING DATE: 15-APR-1998
; APPLICATION NUMBER: US 60/043,553
; FILING DATE: 15-APR-1997
; APPLICATION NUMBER: US 60/048,740
; FILING DATE: 05-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: B. J. Sadoff
; REGISTRATION NUMBER: 36,663
; REFERENCE/DOCKET NUMBER: 620-81
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)816-4091
; TELEFAX: (703)816-4100
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5166 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 23:
US-09-402-923A-23

Query Match 96.6%; Score 4890; DB 4; Length 5166;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 4915; Conservative 0; Mismatches 5; Indels 2; Gaps 2;

QY 142 GCCTCGCGCTCCTGCTATTGTCACCGCGGACGTCACGCTGCTGTCGACCGCGCA 201

Db 217 GCCTCGCCGCTCCTGCTATTTGCCAAACCGCCGGGACGTACGGCTGGTGGACGCCGGGGGA 276
QY 202 GTCAGCTGGAGTCCACCATCGTGGTCAGCGCCCTGGAGGATCGGCCGCGAGTGACTTC 261
Db 277 GTCAGCTGGAGTCCACCATCGTGGTCAGCGCCCTGGAGGATCGGCCGCGAGTGACTTC 336
QY 262 CAGTTTTTCCAAAGGAGCGCTGTACTGGACAGACGTGACGAGGAGGCCCATCAAGCAGAC 321
Db 337 CAGTTTTTCCAAAGGAGCGCTGTACTGGACAGACGTGACGAGGAGGCCCATCAAGCAGAC 396
QY 322 TACCTGAACCAAGACGGGGCCCGCTGTGACAGAACTGGTGTATCTCGGGCTGGTCTCTCCC 381
Db 397 TACCTGAACCAAGACGGGGCCCGCTGTGACAGAACTGGTGTATCTCGGGCTGGTCTCTCCC 456
QY 382 GACGSCCTCGCTCGACTCGGGTGGGCAAGAGCTGTACTGGACGGACTCAGAGACCAAC 441
Db 457 GACGSCCTCGCTCGACTCGGGTGGGCAAGAGCTGTACTGGACGGACTCAGAGACCAAC 516
QY 442 CGCATCGAGGTGGCCAACTCAATGGCACATCCCGGAAGGTGCTCTTCTGGCAGGACCTT 501
Db 517 CGCATCGAGGTGGCCAACTCAATGGCACATCCCGGAAGGTGCTCTTCTGGCAGGACCTT 576
QY 502 GACCAAGCTAGGGCATCGCTTTGGACCCCGCTCAOGGGTACATGTACTGGACAGACTGG 561
Db 577 GACCAAGCTAGGGCATCGCTTTGGACCCCGCTCAOGGGTACATGTACTGGACAGACTGG 636
QY 562 GGTGAGACCCCGGATTTAGCGGGCAGGGATGGATGGCAGCACCAGGAATCATTTGTG 621
Db 637 GGTGAGACCCCGGATTTAGCGGGCAGGGATGGATGGCAGCACCAGGAATCATTTGTG 696
QY 622 GACTCGGACATTTACTGGCCCAATGGACTGACCATCGACCTGGAGGACAGAGCTCTAC 681
Db 697 GACTCGGACATTTACTGGCCCAATGGACTGACCATCGACCTGGAGGACAGAGCTCTAC 756
QY 682 TGGGCTGACGCGCAAGCTCAGCTTTCATCCACCGCTGCCAACTTGGACGGCTCGTTCCGGCAG 741
Db 757 TGGGCTGACGCGCAAGCTCAGCTTTCATCCACCGCTGCCAACTTGGACGGCTCGTTCCGGCAG 816
QY 742 AAGTGGTGGAGGGCAGCTGACGACCCCTTCCGCTTGGACCTGACGCTCTCCGGGACACTCTG 801
Db 817 AAGTGGTGGAGGGCAGCTGACGACCCCTTCCGCTTGGACCTGACGCTCTCCGGGACACTCTG 876
QY 802 TACTGGACAGACTGGCAGCCGCTTCATTCATGCTGCAACAGCGCACTGGGGGAG 861
Db 877 TACTGGACAGACTGGCAGCCGCTTCATTCATGCTGCAACAGCGCACTGGGGGAG 936
QY 862 AGGAAGGAGATCCTGAGTGCCTCTACTCACCCATGGACATCCAGGTGCTGAGCCAGGAG 921
Db 937 AGGAAGGAGATCCTGAGTGCCTCTACTCACCCATGGACATCCAGGTGCTGAGCCAGGAG 996
QY 922 CGGACGCTTTCTTCCACACTGCTGTGAGGAGCAATGGCGCTGCTCCACCTGTGC 981
Db 997 CGGACGCTTTCTTCCACACTGCTGTGAGGAGCAATGGCGCTGCTCCACCTGTGC 1056
QY 982 CTGCTGTCCCAGAGGAGCTTTCTACACATGCGCTGCCCCACGGGTGTCAGCTGCAG 1041
Db 1057 CTGCTGTCCCAGAGGAGCTTTCTACACATGCGCTGCCCCACGGGTGTCAGCTGCAG 1116
QY 1042 GACAAAGGAGAGCTGTGAGGAGGAGCCGAGGAGGTGCTGCTGTGCGCCCGGCGGAGC 1101
Db 1117 GACAAAGGAGAGCTGTGAGGAGGAGCCGAGGAGGTGCTGCTGTGCGCCCGGCGGAGC 1176
QY 1102 GACCTAGGAGATCTCGCTGGACACGCGGAGCTTACCGGACATGCTGCTGAGGTGGAC 1161
Db 1177 GACCTAGGAGATCTCGCTGGACACGCGGAGCTTACCGGACATGCTGCTGAGGTGGAC 1236
QY 1162 GACATCGGAGAGCTGCTGAGGAGGAGCTGAGTACGACCGCTAGAGGGCTATGCTACTGGACA 1221
Db 1237 GACATCGGAGAGCTGCTGAGGAGGAGCTGAGTACGACCGCTAGAGGGCTATGCTACTGGACA 1296
QY 1222 GATGACAGAGTTCGGGGCCATCCGACAGGGCGTACCTGGACGGGTCTGGGGCGCAGACGCTG 1281

Db 1297 GATGACAGAGTTCGGGGCCATCCGACAGGGCGTACTCGGACGGGTCTGGGGCGCAGACGCTG 1356
QY 1282 GTCAAACCGAGATCAACGACCCCGATGGGATCGGGTCTGACTGGGTGGCGCGAAACCTC 1341
Db 1357 GTCAAACCGAGATCAACGACCCCGATGGGATCGGGTCTGACTGGGTGGCGCGAAACCTC 1416
QY 1342 TACTGGACCCGACACGCGGACCGCATCGAGGTGACGCGCCTCAACGGCAGCTCCCGC 1401
Db 1417 TACTGGACCCGACACGCGGACCGCATCGAGGTGACGCGCCTCAACGGCAGCTCCCGC 1476
QY 1402 AAGATCTGGTGTGCGAGGACCTTGGACGAGCCCGAGGCATCGACTGCAACCCCGTATG 1461
Db 1477 AAGATCTGGTGTGCGAGGACCTTGGACGAGCCCGAGGCATCGCACTGCAACCCCGTATG 1536
QY 1462 GGCCTCATGTACTGACAGACTGGGAGAGAACCTTAAATCGAGTGTGCCAATTTGAT 1521
Db 1537 GGCCTCATGTACTGACAGACTGGGAGAGAACCTTAAATCGAGTGTGCCAATTTGAT 1596
QY 1522 GGGCAGGAGCGCGCTGTGCTGGTCAATGCTCTCCCTCGGGTGGCCCAACGGCGTGGCCCTG 1581
Db 1597 GGGCAGGAGCGCGCTGTGCTGGTCAATGCTCTCCCTCGGGTGGCCCAACGGCGTGGCCCTG 1656
QY 1582 GACCTGAGGAGGGGAAAGCTCTACTGGGAGACGCGCAAGACAGACATCGAGGTATC 1641
Db 1657 GACCTGAGGAGGGGAAAGCTCTACTGGGAGACGCGCAAGACAGACATCGAGGTATC 1716
QY 1642 AATGTTGATGGGACGAGAGCGGACCTCTCTGGAGGACAAAGTCCCGCACATTTTCGGG 1701
Db 1717 AATGTTGATGGGACGAGAGCGGACCTCTCTGGAGGACAAAGTCCCGCACATTTTCGGG 1776
QY 1702 TTCAAGCTGTGGGGGACTTTCATCTACTGGACTGACTGGCAGCGCGCAGCATCGAGGG 1761
Db 1777 TTCAAGCTGTGGGGGACTTTCATCTACTGGACTGACTGGCAGCGCGCAGCATCGAGGG 1836
QY 1762 GTGCAAGAGTCAAGGCGCAGCGGAGAGCTCATATGACAGCTGCGCCGACTGATGGG 1821
Db 1837 GTGCAAGAGTCAAGGCGCAGCGGAGAGCTCATATGACAGCTGCGCCGACTGATGGG 1896
QY 1822 CTCAAAGCTGTAATGTGGCCAAAGTCTCGGAAACCAACCGTGTGCGGACAGGAACGG 1881
Db 1897 CTCAAAGCTGTAATGTGGCCAAAGTCTCGGAAACCAACCGTGTGCGGACAGGAACGG 1956
QY 1882 GGGTGCAGCCACCTGTGCTTTCACACCCCAACCGCGTGTGGCTGCGCCCATCGGC 1941
Db 1957 GGGTGCAGCCACCTGTGCTTTCACACCCCAACCGCGTGTGGCTGCGCCCATCGGC 2016
QY 1942 CTGAGCTGTGATGATGACAGACCTGCAATCGTGTGCTGAGGCGCTTCTGGTCTTAC 2001
Db 2017 CTGAGCTGTGATGATGACAGACCTGCAATCGTGTGAGGCGCTTCTGGTCTTAC 2075
QY 2002 AGCAGCGCGCATCCACAGGATCTCCCTCGAGACCAATACAGCGTGGCCATCCG 2061
Db 2076 AGCAGCGCGCATCCACAGGATCTCCCTCGAGACCAATACAGCGTGGCCATCCG 2135
QY 2062 CTCAGCGCGCTCAAGAGGCGCTCAGCCCTGACCTTTGATGTGTCACAAACCAATCTAC 2121
Db 2136 CTCAGCGCGCTCAAGAGGCGCTCAGCCCTGACCTTTGATGTGTCACAAACCAATCTAC 2195
QY 2122 TGGACAGAGTCAAGCTGAGAGCCTCAGCGCGCTTTCATGAACGGAGCTCGGTGAG 2181
Db 2196 TGGACAGAGTCAAGCTGAGAGCCTCAGCGCGCTTTCATGAACGGAGCTCGGTGAG 2255
QY 2182 CACGTGTGGAGTTTGGCTTGAATACCCCGAGGGCATGGCGCTGACTGGATGGGCAAG 2241
Db 2256 CACGTGTGGAGTTTGGCTTGAATACCCCGAGGGCATGGCGCTGACTGGATGGGCAAG 2315
QY 2242 AACCTCTACTGGGCGCACCTGGGACCAACAGAAATCGAAGTGGCGCGCTCGACCGGAG 2301
Db 2316 AACCTCTACTGGGCGCACCTGGGACCAACAGAAATCGAAGTGGCGCGCTCGACCGGAG 2375
QY 2302 TTCCGGCAAGTCTCTGTTGGAGGAGCTTGGACAAACCGAGGTGCTGGCGCTGATCCC 2361
Db 2376 TTCCGGCAAGTCTCTGTTGGAGGAGCTTGGACAAACCGAGGTGCTGGCGCTGATCCC 2435

QY	2362	ACCAAGGGCTACATCTACTGGACCGAGTGGGGCGGCAAGCCGAGGATCGTGGGGCCTTC	2421
Db	2436	ACCAAGGGCTACATCTACTGGACCGAGTGGGGCGGCAAGCCGAGGATCGTGGGGCCTTC	2495
QY	2422	ATGGACGGGACCAACTGTCATGACGCTGTGGACAAGGTGGGCGGGCCAAAGCACTCAC	2481
Db	2496	ATGGACGGGACCAACTGTCATGACGCTGTGGACAAGGTGGGCGGGCCAAAGCACTCAC	2555
QY	2482	ATTGACTACGCTGACCGGCTCTACTGGACCGGCTGGGACCAACATGATGAGTCG	2541
Db	2556	ATTGACTACGCTGACCGGCTCTACTGGACCGGCTGGGACCAACATGATGAGTCG	2615
QY	2542	TCCACATGCTGGGTGACGAGGGGTGCTGATGTCGAGAGATCTCCCGACCCGTTCCGT	2601
Db	2616	TCCACATGCTGGGTGACGAGGGGTGCTGATGTCGAGAGATCTCCCGACCCGTTCCGT	2675
QY	2602	CTGACGAGTACAGCGATTATATCTACTGGACAGACTGGAATCTGCACAGCATTTGAGCGG	2661
Db	2676	CTGACGAGTACAGCGATTATATCTACTGGACAGACTGGAATCTGCACAGCATTTGAGCGG	2735
QY	2662	GCCGACAAAGCTAGCGGCGGAAACCGCACTCTCATFCCAGGGCCACCTGGACTTCGTGATG	2721
Db	2736	GCCGACAAAGCTAGCGGCGGAAACCGCACTCTCATFCCAGGGCCACCTGGACTTCGTGATG	2795
QY	2722	GACATCCTGGTGTCCACTCTCCCGCCAGGATGGCCTCAATGACTGTATGCACAAAC	2781
Db	2796	GACATCCTGGTGTCCACTCTCCCGCCAGGATGGCCTCAATGACTGTATGCACAAAC	2855
QY	2782	GGGAGTGTGGGAGCTGTGCTTGGCATCCCGGGCGCCACCGCTGGGGTGGCGCTCA	2841
Db	2856	GGGAGTGTGGGAGCTGTGCTTGGCATCCCGGGCGCCACCGCTGGGGTGGCGCTCA	2915
QY	2842	CATTACACCTTGAGCCCGCAGCGGCAACTGCAAGCCCGCCCAACCACTTCTTGCTGTC	2901
Db	2916	CATTACACCTTGAGCCCGCAGCGGCAACTGCAAGCCCGCCCAACCACTTCTTGCTGTC	2975
QY	2902	AGCCAGAAATCTGCCATCAGTTCGGATGATCCCGGACGACGACGACGCGGATCTCATC	2961
Db	2976	AGCCAGAAATCTGCCATCAGTTCGGATGATCCCGGACGACGACGACGCGGATCTCATC	3035
QY	2962	CTGCCCCTGCACTGAGGAAAGTCAAGCCATCGACTATGACCCACTGGAAGTTTC	3021
Db	3036	CTGCCCCTGCACTGAGGAAAGTCAAGCCATCGACTATGACCCACTGGAAGTTTC	3095
QY	3022	ATCTACTGGGTGATGGCGCCAGAAATCAAGGAGGCAAGGAGCGAGCGGACCCAGCCC	3081
Db	3096	ATCTACTGGGTGATGGCGCCAGAAATCAAGGAGGCAAGGAGCGAGCGGACCCAGCCC	3155
QY	3082	TTTTGTTTTGACCTCTCTGAGCCAGGCCAAACCCAGACGAGCGAGCCCAAGCTCAGC	3141
Db	3156	TTTTGTTTTGACCTCTCTGAGCCAGGCCAAACCCAGACGAGCGAGCCCAAGCTCAGC	3215
QY	3142	ATCGACATCTACGCGGACACTGTTCTGGACGTCGAGGCGCACCAATACCATCAACGTC	3201
Db	3216	ATCGACATCTACGCGGACACTGTTCTGGACGTCGAGGCGCACCAATACCATCAACGTC	3275
QY	3202	CACAGGCTGAGCGGGGAAGCCATGGGGGTGTGTGCTGTGGGAGCCGCGCAACAGCCAGG	3261
Db	3276	CACAGGCTGAGCGGGGAAGCCATGGGGGTGTGTGCTGTGGGAGCCGCGCAACAGCCAGG	3335
QY	3262	GCCATCGTCTCAAGCGGGAGGGGTACTGTACTTCCACCAATGCAGGACCGGGCA	3321
Db	3336	GCCATCGTCTCAAGCGGGAGGGGTACTGTACTTCCACCAATGCAGGACCGGGCA	3395
QY	3322	GCCAAAGATCGAAACCGGACCGCTTGGACCGGACCCGAGCGAGGTCTTTCACCAACCGGC	3381
Db	3396	GCCAAAGATCGAAACCGGACCGCTTGGACCGGACCCGAGCGAGGTCTTTCACCAACCGGC	3455
QY	3382	CTCATTCGCGCCCTGTGGCCCTGTGGTGTGGAACAACAATGGGGCAAGCTGTTCTGGGTGAC	3441
Db	3456	CTCATTCGCGCCCTGTGGCCCTGTGGTGTGGAACAACAATGGGGCAAGCTGTTCTGGGTGAC	3515

QY	3442	GCGGACCTGAGCGCATTTGAGAGCTGTACCTGTGAGGGGCAACCGCCTGACCCCTGGAG	3501
Db	3516	GCGGACCTGAGCGCATTTGAGAGCTGTGACCTGTGAGGGGCAACCGCCTGACCCCTGGAG	3575
QY	3502	GACGCCAAACATCTGTGACAGCTCTGGGCTGTGACCATCTTGGCAAGCATCTCTACTGGATC	3561
Db	3576	GACGCCAAACATCTGTGACAGCTCTGGGCTGTGACCATCTTGGCAAGCATCTCTACTGGATC	3635
QY	3562	GACGCCAGAGCAGATGATGAGCGTGTGAGAGAACCAACCGGGGAACAAGCGGACTCGC	3621
Db	3636	GACGCCAGAGCAGATGATGAGCGTGTGAGAGAACCAACCGGGGAACAAGCGGACTCGC	3695
QY	3622	ATCCAGGGCCGTGTGCGCCACTCTCACTGGGCATCCATGCAAGTGGAGGAGTCAAGCTGGAG	3681
Db	3696	ATCCAGGGCCGTGTGCGCCACTCTCACTGGGCATCCATGCAAGTGGAGGAGTCAAGCTGGAG	3755
QY	3682	GAGTTCTCAGCCCAACCCATGTGCGCGTGCACAAATGGTGGCTGCTCCCACTCTGTTATGCC	3741
Db	3756	GAGTTCTCAGCCCAACCCATGTGCGCGTGCACAAATGGTGGCTGCTCCCACTCTGTTATGCC	3815
QY	3742	AAGGGTATGGGAACAACAAGTGTCTATGCCAGTCCACCTCTGCTGCTCTGAGAACCTG	3801
Db	3816	AAGGGTATGGGAACAACAAGTGTCTATGCCAGTCCACCTCTGCTGCTCTGAGAACCTG	3875
QY	3802	CTGACCTGTGGAGAGCCGCCACCTGCTCCCGGACCAAGTTTGCATGTGCACAGGGGAG	3861
Db	3876	CTGACCTGTGGAGAGCCGCCACCTGCTCCCGGACCAAGTTTGCATGTGCACAGGGGAG	3935
QY	3862	ATCGACTGTATCCCGGGGGCTGGCGCTGTGACGGCTTTCCCGAGTGCATGACACAGAGC	3921
Db	3936	ATCGACTGTATCCCGGGGGCTGGCGCTGTGACGGCTTTCCCGAGTGCATGACACAGAGC	3995
QY	3922	GACGAGAGGGCTGCGCGCTGTGCTCCGCGCCAGTTCCCTCTGGCGCGGGGGTCAGTGT	3981
Db	3996	GACGAGAGGGCTGCGCGCTGTGCTCCGCGCCAGTTCCCTCTGGCGCGGGGGTCAGTGT	4055
QY	3982	GTGACCTGTGCGCTGCGCGAGGAGGACAGACTGTGAGGACCGCTCAGACAGAGGCG	4041
Db	4056	GTGACCTGTGCGCTGCGCGAGGAGGACAGACTGTGAGGACCGCTCAGACAGAGGCG	4114
QY	4042	GACTGTGACGCGCATCTGCTGCCCAACAGTTCGGTGTGCGAGCGGCGAGTGTGCTCCTC	4101
Db	4116	GACTGTGACGCGCATCTGCTGCCCAACAGTTCGGTGTGCGAGCGGCGAGTGTGCTCCTC	4174
QY	4102	ATCAAAACAGAGTGGACTCTTCCCGACTGTATCGACGCGCTCCGAGAGTCAATGTGT	4161
Db	4176	ATCAAAACAGAGTGGACTCTTCCCGACTGTATCGACGCGCTCCGAGAGTCAATGTGT	4234
QY	4162	GAAATCACCAGCGCCCTCAGACGACAGCCCGGCCCAACAGCAGTGCATCGGGCCCGTC	4221
Db	4236	GAAATCACCAGCGCCCTCAGACGACAGCCCGGCCCAACAGCAGTGCATCGGGCCCGTC	4294
QY	4222	ATTGGCATCATCTCTCTCTCTGTCATGGGTGGTGTCTATTTTGTGTGTCAGCGCGTG	4281
Db	4296	ATTGGCATCATCTCTCTCTCTGTCATGGGTGGTGTCTATTTTGTGTGTCAGCGCGTG	4354
QY	4282	GTGTGCCAGCGCTATGCGGGGGCCAAACCGGGCCCTTCCGCGACGAGTATGTCAAGCGGAC	4341
Db	4356	GTGTGCCAGCGCTATGCGGGGGCCAAACCGGGCCCTTCCGCGACGAGTATGTCAAGCGGAC	4414
QY	4342	CCGACAGTGGCCCTCAATTTTCATAGCCCGGGGGTTTCCAGCATGGCCCTTTCACAGGC	4401
Db	4416	CCGACAGTGGCCCTCAATTTTCATAGCCCGGGGGTTTCCAGCATGGCCCTTTCACAGGC	4474
QY	4402	ATCGCATCGGAAAGTCCATGATGAGCTCGTGTAGCGCTGTATGGGGGGCCGGGGGGGTG	4461
Db	4476	ATCGCATCGGAAAGTCCATGATGAGCTCGTGTAGCGCTGTATGGGGGGCCGGGGGGGTG	4534
QY	4462	CCCTGTAGCAGCCGGAACAGCTCAAGGGGCTCTGTCAGAGCTGTGTCAGCAGCAGAG	4521
Db	4536	CCCTGTAGCAGCCGGAACAGCTCAAGGGGCTCTGTCAGAGCTGTGTCAGCAGCAGAG	4594
QY	4522	GCCACGCTGTACCGCGCGATCTCTGAAACCGCGCGCCCTCTCCCGGCGCACGGAACCTCTCCTG	4581

[illegible]

Search completed: February 18, 2005, 20:21:58
Job time : 811 secs

THIS PAGE BLANK (USPTO)

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	5061.4	100.0	5135	6	AX658229	Sequence
2	5061.4	100.0	5135	9	AF077820	Homo sapi
3	5058.2	99.9	5120	6	AR567475	Sequence
4	5058.2	99.9	5120	6	AX277522	Sequence
5	5058.2	99.9	5120	6	AX118086	Sequence
6	5056.8	99.9	5099	6	CQ870240	Sequence
7	5056.6	99.9	5120	6	AR567476	Sequence
8	5056.6	99.9	5120	6	AX277523	Sequence
9	5056.6	99.9	5120	6	AX118087	Sequence
10	5052	99.8	5100	6	AX821900	Sequence
11	5052	99.8	5100	9	AF064548	Homo sapi
12	5028	99.3	5098	6	AR305075	Sequence
13	5028	99.3	5098	6	AR309179	Sequence
14	5028	99.3	5098	6	BD105986	Novel LDL
15	5010.8	99.0	5135	6	CQ870242	Sequence
16	4999	98.7	5059	9	AB017498	Homo sapi
17	4902	96.8	5022	6	AR305086	Sequence
18	4902	96.8	5022	6	AR309190	Sequence
19	4902	96.8	5022	6	BD105997	Novel LDL

Db 218 CCGCTGTGGAGCGCGCGGAGTCAAGCTGGAGTCCACATCATGTGGTTCAGCGGCTGGAG 277
Qy 241 GATCGGCGCAGTGGAGCTTCCAGTCTTCCAAAGGAGCGGTGTACTGTGACAGAGCTGAGC 300
Db 278 GATCGGCGCAGTGGAGCTTCCAGTCTTCCAAAGGAGCGGTGTACTGTGACAGAGCTGAGC 337
Qy 301 GAGGAGGCGCATCAAGCAGACCTACTGAAACAGACGCGGCGCGCTGTGACAGAACTGGTC 360
Db 338 GAGGAGGCGCATCAAGCAGACCTACTGAAACAGACGCGGCGCGCTGTGACAGAACTGGTC 397
Qy 361 ATCTCCGCGCTGGTCTCTCCGACGCGCTCGCTCGGACTGGGTGGGCGGAGAGCTGTAC 420
Db 398 ATCTCCGCGCTGGTCTCTCCGACGCGCTCGCTCGGACTGGGTGGGCGGAGAGCTGTAC 457
Qy 421 TGGACGAGCTCAGAGCAACCGCATCGAGTGGCCAACTCAATGGCACAATCCCGGAAG 480
Db 458 TGGACGAGCTCAGAGCAACCGCATCGAGTGGCCAACTCAATGGCACAATCCCGGAAG 517
Qy 481 GTGCTCTTCTGGCAGGACCTTTGACAGCCTTAGGGCCATCGCCTTGGACCCCGCTCAAGG 540
Db 518 GTGCTCTTCTGGCAGGACCTTTGACAGCCTTAGGGCCATCGCCTTGGACCCCGCTCAAGG 577
Qy 541 TACATGTACTGACAGACTGGGTGAGACGCGCGGATTTGACGCGGCAAGGATGATGGC 600
Db 578 TACATGTACTGACAGACTGGGTGAGACGCGCGGATTTGACGCGGCAAGGATGATGGC 637
Qy 601 AGCACCCGGAAGATCAATGTGAGCTCGGACATTTACTGGCCCAATGAGACTGACCATCGAC 660
Db 638 AGCACCCGGAAGATCAATGTGAGCTCGGACATTTACTGGCCCAATGAGACTGACCATCGAC 697
Qy 661 CTGGAGGACGAGACTCTACTTGGGCTGACGCGCAAGCTCAGCTTCAATCGACGCTGCCAAC 720
Db 698 CTGGAGGACGAGACTCTACTTGGGCTGACGCGCAAGCTCAGCTTCAATCGACGCTGCCAAC 757
Qy 721 CTGGAGGCTCTTCCGAGAGGCTGGTGGAGGCGAGCTGACGACACCCCTTGGCCCTG 780
Db 758 CTGGAGGCTCTTCCGAGAGGCTGGTGGAGGCGAGCTGACGACACCCCTTGGCCCTG 817
Qy 781 AGCGCTCTCCGGGACACTCTGTACTGACAGACTGGCAGACCCGCTCCATCCATCCCTGC 840
Db 818 AGCGCTCTCCGGGACACTCTGTACTGACAGACTGGCAGACCCGCTCCATCCATCCCTGC 877
Qy 841 AACAGCGCACTGGGGGAGAGGAGAGATCTGTAGTGGCTTACTTCAACCATGAC 900
Db 878 AACAGCGCACTGGGGGAGAGGAGAGATCTGTAGTGGCTTACTTCAACCATGAGC 937
Qy 901 ATCCAGGTCTGAGCAGAGCGGAGCCTTCTTCCACACTCGCTGTGAGGAGGACAAAT 960
Db 938 ATCCAGGTCTGAGCAGAGCGGAGCCTTCTTCCACACTCGCTGTGAGGAGGACAAAT 997
Qy 961 GCGGCTGTCCCACTGTGCTGTGTCTCCCAAGCGAGCCTTTTACATGTGCGCTGC 1020
Db 998 GCGGCTGTCTCCCACTGTGCTGTGTCTCCCAAGCGAGCCTTTTACACATGCGCCTGC 1057
Qy 1021 CCCACGGGTGTGAGCTGAGCAGACGAGAGAGCTGTAAAGCAGAGCGCGGAGAGGTG 1080
Db 1058 CCCACGGGTGTGAGCTGAGCAGACGAGAGAGCTGTAAAGCAGAGCGCGGAGAGGTG 1117
Qy 1081 CTGCTCTGTGCGCGGAGGAGCTTACGAGAGGATCTCGCTGACACGCGGAGCTTCAAC 1140
Db 1118 CTGCTCTGTGCGCGGAGGAGCTTACGAGAGGATCTCGCTGACACGCGGAGCTTCAAC 1177
Qy 1141 GACATCGTGTGAGTGGAGAGCATCTCGGACGCGCATTTGGCATCGACTACGACCGGTA 1200
Db 1178 GACATCGTGTGAGTGGAGAGCATCTCGGACGCGCATTTGGCATCGACTACGACCGGTA 1237
Qy 1201 GAGGCTATGTCTTACTGACAGATGAGGTGCGGCGCATCCGAGGCGGTACTTGGAC 1260
Db 1238 GAGGCTATGTCTTACTGACAGATGAGGTGCGGCGCATCCGAGGCGGTACTTGGAC 1297
Qy 1261 GGGTCTGGGCGGAGAGCTGTGTCACACCGAGATCAAGACCCCGATGCGCGTGC 1320

Db 1298 GGGTCTGGGCGGAGAGCTGTGTCACACCGAGATCAACGACCCCGATGCGCGTGC 1357
Qy 1321 GACTGGGTGGCGCGGAAACCTCTTACTGGAACGACACGCGGACCGCATCGAGGTGACG 1380
Db 1358 GACTGGGTGGCGCGGAAACCTCTTACTGGAACGACACGCGGACCGCATCGAGGTGACG 1417
Qy 1381 CGCTCAACGCGACCTTCCGCAAGATCTTGGTGTCTGGAGGACCTTGACGAGCCCGGAGCC 1440
Db 1418 CGCTCAACGCGACCTTCCGCAAGATCTTGGTGTCTGGAGGACCTTGACGAGCCCGGAGCC 1477
Qy 1441 ATCGCACTGACACCCCGTGTATGGGCTCATGTACTGGAACAGACTCTGGGAGAGAACCTTAAA 1500
Db 1478 ATCGCACTGACACCCCGTGTATGGGCTCATGTACTGGAACAGACTCTGGGAGAGAACCTTAAA 1537
Qy 1501 ATCGAGTGTCCAACTTGGATGGGAGGAGCGGCTGTGTGTCTCAATGCTCCCTCGGG 1560
Db 1538 ATCGAGTGTCCAACTTGGATGGGAGGAGCGGCTGTGTGTCTCAATGCTCCCTCGGG 1597
Qy 1561 TGGCCCAACCGCTTGGGCTTGGACCTTGACGAGAGGGAAGCTCTACTGGGAGAGCGCAAG 1620
Db 1598 TGGCCCAACCGCTTGGGCTTGGACCTTGACGAGAGGGAAGCTCTACTGGGAGAGCGCAAG 1657
Qy 1621 ACAGACAAGATCGAGGTGATCAATGTGTGATGGAGAGAGGCGGACCTCTCTGGAGGAC 1680
Db 1658 ACAGACAAGATCGAGGTGATCAATGTGTGATGGAGAGAGGCGGACCTCTCTGGAGGAC 1717
Qy 1681 AAGCTCCGCGACATTTTGGGTTTCCAGCTCTGGGGACTTCACTACTTGAGTGAATG 1740
Db 1718 AAGCTCCGCGACATTTTGGGTTTCCAGCTCTGGGGACTTCACTACTTGAGTGAATG 1777
Qy 1741 CAGCGCGCAGCATCGAGCGGCTGCAACAGGTCAAGGCGCAGCGGAGCGTCAATCATTTGAC 1800
Db 1778 CAGCGCGCAGCATCGAGCGGCTGCAACAGGTCAAGGCGCAGCGGAGCGTCAATCATTTGAC 1837
Qy 1801 CAGCTGCCGACCTGTATGGGCTCAAAAGCTGTGAATGTGGCCAGGTCTGGAGACCAAC 1860
Db 1838 CAGCTGCCGACCTGTATGGGCTCAAAAGCTGTGAATGTGGCCAGGTCTGGAGACCAAC 1897
Qy 1861 CCGTGTGGGACAGGAAACGCGGCTGACGACCTGTGTCTTTCACACCCCGACCAAC 1920
Db 1898 CCGTGTGGGACAGGAAACGCGGCTGACGACCTGTGTCTTTCACACCCCGACCAAC 1957
Qy 1921 CCGTGTGGCTGCCCATCGGCTCGAGCTGTGATGATGACATGAAAGACCTGTGATCGTGCCT 1980
Db 1958 CCGTGTGGCTGCCCATCGGCTCGAGCTGTGATGATGACATGAAAGACCTGTGATCGTGCCT 2017
Qy 1981 GAGGCTTCTTGTCTTTCACAGCAGAGCGGCTTCCACAGATCTCCCTCGAGACCAAT 2040
Db 2018 GAGGCTTCTTGTGTCTTTCACAGCAGAGCGGCTTCCACAGATCTCCCTCGAGACCAAT 2077
Qy 2041 AACACGAGTGGGCTATCCGCTCAACGCGCTCAAGGAGGCTCAGCCCTCGACTTTGAT 2100
Db 2078 AACACGAGTGGGCTATCCGCTCAACGCGCTCAAGGAGGCTCAGCCCTCGACTTTGAT 2137
Qy 2101 GTGTCCAAACACCACTTACTGGAACAGCTGAGCTGAAAGACCATCAGCGCGGCTTC 2160
Db 2138 GTGTCCAAACACCACTTACTGGAACAGCTGAGCTGAAAGACCATCAGCGCGGCTTC 2197
Qy 2161 ATGAACGGAGTGTGAGTGGAGCAGTGTGTGAGTTTGGCTTGTACTACCCGAGGCGCATG 2220
Db 2198 ATGAACGGAGTGTGAGTGGAGCAGTGTGTGAGTTTGGCTTGTACTACCCGAGGCGCATG 2257
Qy 2221 GCGGTGTGATGAGTGGGCAAGAACCTTACTGGGCGGACACTGGGACCAAGAAATCGAA 2280
Db 2258 GCGGTGTGATGAGTGGGCAAGAACCTTACTGGGCGGACACTGGGACCAAGAAATCGAA 2317
Qy 2281 GTGGCGGCTTGGACGGGAGTTCCGGCAAGTCTCTGTGTGGAGGACTTGGACAAACCG 2340
Db 2318 GTGGCGGCTTGGACGGGAGTTCCGGCAAGTCTCTGTGTGGAGGACTTGGACAAACCG 2377
Qy 2341 AGGTCTGGCTTGGATCCCAACCAAGGCTTACTACTTGAGACCGAGTGGGCGGCAAG 2400
Db 2378 AGGTCTGGCTTGGATCCCAACCAAGGCTTACTACTTGAGACCGAGTGGGCGGCAAG 2437

Db 578 TACATGTACTGACACAGACTGGGTGAGACGCCCCGGATTGAGCGGCAGGGATGATGGC 637
Qy AGCACCCGGAGAGATCATTTGTGACACTCGACACATTTACTTGGCCCAATGGAATGACCATCGAC 660
Db 638 AGCACCCGGAGAGATCATTTGTGACACTCGACACATTTACTTGGCCCAATGGAATGACCATCGAC 697
Qy 661 CTGAGGAGCAGAAAGCTCTACTGGGCTGACGCCAAGCTCAGCTTCATCCACCGTGCACAC 720
Db 698 CTGAGGAGCAGAAAGCTCTACTGGGCTGACGCCAAGCTCAGCTTCATCCACCGTGCACAC 757
Qy 721 CTGAGCGGCTCTGTCGCGCAGAAAGTGTGTGAGGCGACGCTGACGCAACCCCTTCCGCTTG 780
Db 758 CTGAGCGGCTCTGTCGCGCAGAAAGTGTGTGAGGCGACGCTGACGCAACCCCTTCCGCTTG 817
Qy 781 ACGCTCTCCGGGACACTCTCTACTGAGCAGACTGGCAGACCCGCTCCATCCATCGCTGC 840
Db 818 ACGCTCTCCGGGACACTCTCTACTGAGCAGACTGGCAGACCCGCTCCATCCATCGCTGC 877
Qy 841 AACAGCGCACTGGGGGAGAGAGAGAGATCTCTGAGTGCCTCTACTCAACCCATGGAC 900
Db 878 AACAGCGCACTGGGGGAGAGAGAGAGATCTCTGAGTGCCTCTACTCAACCCATGGAC 937
Qy 901 ATCCAGGTGTGAGCAGAGAGCGGACGCTTTCTTCCACATCTGCTGTGAGGAGGACAA 960
Db 938 ATCCAGGTGTGAGCAGAGAGCGGACGCTTTCTTCCACATCTGCTGTGAGGAGGACAA 997
Qy 961 GCGGCTGCTCCCACTGCTGCTGCTGCTGCCCAAGCGAGCCTTTCTACACATGCGCTGC 1020
Db 998 GCGGCTGCTCCCACTGCTGCTGCTGCCCAAGCGAGCCTTTCTACACATGCGCTGC 1057
Qy 1021 CCCACGGGTGTGAGCTGCAGACAAACGGCAGGACGTTAGGAGAGAGCGGAGAGGTG 1080
Db 1058 CCCACGGGTGTGAGCTGCAGACAAACGGCAGGAGTGTAGGAGAGAGCGGAGAGGTG 1117
Qy 1081 CTGCTGTGCGCCGCGGAGCGGACCTTACGGAGATCTGCTGTGACACGCGGAGCTTCAAC 1140
Db 1118 CTGCTGTGCGCCGCGGAGCGGACCTTACGGAGATCTGCTGTGACACGCGGAGCTTCAAC 1177
Qy 1141 GACATCGTGTGAGGTGAGACGATCCGGCAGCGCATTTGCCATCGATAGCACCCGCTA 1200
Db 1178 GACATCGTGTGAGGTGAGACGATCCGGCAGCGCATTTGCCATCGATAGCACCCGCTA 1237
Qy 1201 GAGGCTATGTCTACTGACAGATGACGAGGTGCGGCGCATCCGACAGCGGTACTCTGCAC 1260
Db 1238 GAGGCTATGTCTACTGACAGATGACGAGGTGCGGCGCATCCGACAGCGGTACTCTGCAC 1297
Qy 1261 GGGTCTGGGGCGCAGACGCTGGTCAACACCGAGATCAACGACCCCGATGCGATCGCGTC 1320
Db 1298 GGGTCTGGGGCGCAGACGCTGGTCAACACCGAGATCAACGACCCCGATGCGATCGCGTC 1357
Qy 1321 GACTGGGTGGCCGAAACCTCTACTGACCGACACGCGGACGACCGGATCGAGGTGACG 1380
Db 1358 GACTGGGTGGCCGAAACCTCTACTGACCGACACGCGGACGACCGGATCGAGGTGACG 1417
Qy 1381 CGCCTCAACGCGACCTCCGCAAGATCTGTGTGTCGAGGACCTGAGACGAGCCCGAGCC 1440
Db 1418 CGCCTCAACGCGACCTCCGCAAGATCTGTGTGTCGAGGACCTGAGACGAGCCCGAGCC 1477
Qy 1441 ATCGCACTGCAACCCGCTGATGGGCTCATGTACTGGAACAGCTGAGGAGAGAACCTTAA 1500
Db 1478 ATCGCACTGCAACCCGCTGATGGGCTCATGTACTGGAACAGCTGAGGAGAGAACCTTAA 1537
Qy 1501 ATCGAGTGTGCAACTTGGATGGGACGAGCGGCTGTGCTGTCATGCTCCCTCGG 1560
Db 1538 ATCGAGTGTGCAACTTGGATGGGACGAGCGGCTGTGCTGTCATGCTCCCTCGG 1597
Qy 1561 TGGCCCAACGCGCTTGGCCCTGAGCTGACGAGGAGAGCTTACTTGGGAGAGCGCAAG 1620
Db 1598 TGGCCCAACGCGCTTGGCCCTGAGCTGACGAGGAGAGAGCTTACTTGGGAGAGCGCAAG 1657
Qy 1621 ACAGACAAGATCGAGGTGATCAATGTTGATGGGACGAGGCGGACCCCTCTGGAGGAC 1680

Db 1658 ACAGACAAGATCGAGGTGATCAATGTTGATGGGACGAGGCGGACCCCTCTCTGGAGAC 1717
Qy 1681 AAGCTCCCGCACATTTTCGGGTTCAAGCTGCTGGGGAGCTTTCATCTACTGACGACTGAC 1740
Db 1718 AAGCTCCCGCACATTTTCGGGTTCAAGCTGCTGGGGAGCTTTCATCTACTGACGACTGAC 1777
Qy 1741 CAGCGCGCAGCATCGAGCGGGTGCAAAAGGTCAAGGCCAGCCGGGAGCTCATCATTTGAC 1800
Db 1778 CAGCGCGCAGCATCGAGCGGGTGCAAAAGGTCAAGGCCAGCCGGGAGCTCATCATTTGAC 1837
Qy 1801 CAGCTGCCCGACCTGATGGGCTCAAAAGCTGTGAATGTGGCCAAAGGTGCTGCGAACCAAC 1860
Db 1838 CAGCTGCCCGACCTGATGGGCTCAAAAGCTGTGAATGTGGCCAAAGGTGCTGCGAACCAAC 1897
Qy 1861 CCGTGTGCGGACAGGACGCGGGGTGAGCCACCTGCTGCTTTCACACCCACCGCAAC 1920
Db 1898 CCGTGTGCGGACAGGACGCGGGGTGAGCCACCTGCTGCTTTCACACCCACCGCAAC 1957
Qy 1921 CCGTGTGCTGCCCATCGGCTTGGAGCTGTGAGTGACATGAAGACCTGCTGCTGCT 1980
Db 1958 CCGTGTGCTGCCCATCGGCTTGGAGCTGTGAGTGACATGAAGACCTGCTGCTGCT 2017
Qy 1981 GAGGCTTTCTTGTCTTCAACAGCAGAGCCGCCATCCACAGATCTCCCTCGAGACCAAT 2040
Db 2018 GAGGCTTTCTTGTCTTCAACAGCAGAGCCGCCATCCACAGATCTCCCTCGAGACCAAT 2077
Qy 2041 AACAAACGCTGGGCATCCGCTCAACGGGCTCAAGGAGGCTCAGCCCTGGAATTTGAT 2100
Db 2078 AACAAACGCTGGGCATCCGCTCAACGGGCTCAAGGAGGCTCAGCCCTGGAATTTGAT 2137
Qy 2101 GTGTCTCAACCAACCATCTACTGACAGACGCTCAGCTTGAAGACCATCAGCCGCGCTTC 2160
Db 2138 GTGTCTCAACCAACCATCTACTGACAGACGCTCAGCTTGAAGACCATCAGCCGCGCTTC 2197
Qy 2161 ATGAAACGCGAGCTCGGTGAGACGCTGTGAGTTTGGCTTGAATACCCCGAGGCGATG 2220
Db 2198 ATGAAACGCGAGCTCGGTGAGACGCTGTGAGTTTGGCTTGAATACCCCGAGGCGATG 2257
Qy 2221 GCGCTTCACTGGAATGGGCAAGACCTCTACTGCGGCGACACTGGGACCAACAGATCCAA 2280
Db 2258 GCGCTTCACTGGAATGGGCAAGACCTCTACTGCGGCGACACTGGGACCAACAGATCCAA 2317
Qy 2281 GTGCGCGGCTGGAACGCGGAGTTCCGCAAGTCTCGTGTGAGGAGCTTGGACAAACCGG 2340
Db 2318 GTGCGCGGCTGGAACGCGGAGTTCCGCAAGTCTCGTGTGAGGAGCTTGGACAAACCGG 2377
Qy 2341 AGGTGCTGCGCTTGAATCCCAACAAAGGCTTACATCTAGGACGAGTGGGGGCGCAAG 2400
Db 2378 AGGTGCTGCGCTTGAATCCCAACAAAGGCTTACATCTAGGACGAGTGGGGGCGCAAG 2437
Qy 2401 CCGAGGATCGTGGGCTTTCATGGAACGAGCACTGCACTGAGCGCTGCTGGAACAGGTG 2460
Db 2438 CCGAGGATCGTGGGCTTTCATGGAACGAGCACTGCACTGAGCGCTGCTGGAACAGGTG 2497
Qy 2461 GCGCGGCGCAACGACCTTCACTTGAATGAGCTGACCGAGCGCTTACTTGAACCGACCTG 2520
Db 2498 GCGCGGCGCAACGACCTTCACTTGAATGAGCTGACCGAGCGCTTACTTGAACCGACCTG 2557
Qy 2521 GACACCAACATGATCGAGTGTGCTGCAACATGCTGGGTGAGGAGCGGCTGATGTCGAC 2580
Db 2558 GACACCAACATGATCGAGTGTGCTGCAACATGCTGGGTGAGGAGCGGCTGATGTCGAC 2617
Qy 2581 GATCTCCGCGACCGGCTTTCGAGCGAGTGTGAGCGATTTATATCTACTTGAACAGCTG 2640
Db 2618 GATCTCCGCGACCGGCTTTCGAGCGAGTGTGAGCGATTTATATCTACTTGAACAGCTG 2677
Qy 2641 AATCTGACAGCATTTGAGCGGCGCAACAGATAGCGGCGGAAACCGCACCTCATTCAG 2700
Db 2678 AATCTGACAGCATTTGAGCGGCGCAACAGATAGCGGCGGAAACCGCACCTCATTCAG 2737
Qy 2701 GGCCACCTGAGCTTCTGATGAGACATCTCTGGTGTTCATCTCTCCCGCAGGATGGCTC 2760
Db 2738 GGCCACCTGAGCTTCTGATGAGACATCTCTGGTGTTCATCTCTCCCGCAGGATGGCTC 2797

Db 1640 ACAGACAAAGATCGAGGTGATCAATGTTGATGGACGAAGAGCGGACCCCTCTCTGGAGGAC 1699
QY 1681 AAGCTCCCGCACATTTTCGGGTTTACGCTGTCTGGGGGACTTCACTACTGAGACTGACTGG 1740
Db 1700 AAGCTCCCGCACATTTTCGGGTTTACGCTGTCTGGGGGACTTCACTACTGAGACTGACTGG 1759
QY 1741 CAGCGCCGACGATCGAGCGGGTGCAAAAGGTCAAGGCGAGCGGGAGGTCACTCATTTGAC 1800
Db 1760 CAGCGCCGACGATCGAGCGGGTGCAAAAGGTCAAGGCGAGCGGGAGGTCACTCATTTGAC 1819
QY 1801 CAGCTGCCGACCTGATGGGCTCAAAGCTGTGAATGTGGCCAAAGGTCTCGGACCAAC 1860
Db 1820 CAGCTGCCGACCTGATGGGCTCAAAGCTGTGAATGTGGCCAAAGGTCTCGGACCAAC 1879
QY 1861 CCGTGTGGGACAGGAAACGGGGGTGCAGCCACCTGTGCTTCTTTCACACCCCAACGCAACC 1920
Db 1880 CCGTGTGGGACAGGAAACGGGGGTGCAGCCACCTGTGCTTCTTTCACACCCCAACGCAACC 1939
QY 1921 CCGTGTGGCTGCCCATCGGCTTGAGCTGTGAGTGATGACATGAAGACCTGCACTGTGCT 1980
Db 1940 CCGTGTGGCTGCCCATCGGCTTGAGCTGTGAGTGATGACATGAAGACCTGCACTGTGCT 1999
QY 1981 GAGGCTTCTTGGTCTTCAACAGCAGAGCCGCCATCCACAGGATCTCCCTCGAGACCAAT 2040
Db 2000 GAGGCTTCTTGGTCTTCAACAGCAGAGCCGCCATCCACAGGATCTCCCTCGAGACCAAT 2059
QY 2041 AACACGACGTGGCCATCCCGCTCACGGGCGTCAAGGAGGCTCAGCCCTGGACTTTGAT 2100
Db 2060 AACACGACGTGGCCATCCCGCTCACGGGCGTCAAGGAGGCTCAGCCCTGGACTTTGAT 2119
QY 2101 GTGTCCAAACCAACATCTACTGACAGACGCTCAGCCTGAAGACCATCAGCGCGGCTTC 2160
Db 2120 GTGTCCAAACCAACATCTACTGACAGACGCTCAGCCTGAAGACCATCAGCGCGGCTTC 2179
QY 2161 ATGAACGGGAGCTCGTGGAGACGCTGGGCAAGTCTCGTGTGGAGGACTTGAACCCGAGGGCATG 2220
Db 2180 ATGAACGGGAGCTCGTGGAGACGCTGGGCAAGTCTCGTGTGGAGGACTTGAACCCGAGGGCATG 2239
QY 2221 GCCGTGACTGGATGGGCAAGACCTCTACTTGGGCGGACACTGGGACCAACAGATCGAA 2280
Db 2240 GCCGTGACTGGATGGGCAAGACCTCTACTTGGGCGGACACTGGGACCAACAGATCGAA 2299
QY 2281 GTGGCGGCTGGAGCGGCGAGTTCGGCAAGTCTCGTGTGGAGGACTTGAACAAACCGG 2340
Db 2300 GTGGCGGCTGGAGCGGCGAGTTCGGCAAGTCTCGTGTGGAGGACTTGAACAAACCGG 2359
QY 2341 AGGTGCTGGGCTGGATGCCAACAAGGGCTACATCTACTGGAACGAGTGGGGGGCAAG 2400
Db 2360 AGGTGCTGGGCTGGATGCCAACAAGGGCTACATCTACTGGAACGAGTGGGGGGCAAG 2419
QY 2401 CCGAGGATCGTGGGCGCTTCATGGACGGGACCAACTGCATGACGCTGGTGGACAAAGTG 2460
Db 2420 CCGAGGATCGTGGGCGCTTCATGGACGGGACCAACTGCATGACGCTGGTGGACAAAGTG 2479
QY 2461 GCGCGGGCAACGACCTCACCATTTGACTACGCTGACAGCGCTTACTTGGACCGACCTG 2520
Db 2480 GCGCGGGCAACGACCTCACCATTTGACTACGCTGACAGCGCTTACTTGGACCGACCTG 2539
QY 2521 GACACCAACATGATGAGTGTGCCAACAATGTGGGTCAAGGACGGGTCTGATTCGCGAC 2580
Db 2540 GACACCAACATGATGAGTGTGCCAACAATGTGGGTCAAGGACGGGTCTGATTCGCGAC 2599
QY 2581 GATCTCCGCGACCGCTTCGCTGACGAGTACGAGGATATATCTACTGACAGACTGG 2640
Db 2600 GATCTCCGCGACCGCTTCGCTGACGAGTACGAGGATATATCTACTGACAGACTGG 2659
QY 2641 AATCTGCAACGATTTGAGCGGGCGGACAAAGACTAGCGGCGGAAACCGCACCTCATCCAG 2700
Db 2660 AATCTGCAACGATTTGAGCGGGCGGACAAAGACTAGCGGCGGAAACCGCACCTCATCCAG 2719
QY 2701 GGCACCTTGGATTCGTGATGACATCTCTGGTGTTCACCTCTCCGCGCAGGATGGCCTC 2760
Db 2720 GGCACCTTGGATTCGTGATGACATCTCTGGTGTTCACCTCTCCGCGCAGGATGGCCTC 2779

QY 2761 AATGACTGTATGCACACAAACGGGCGAGTGTGGGCGAGTGTGCCTTGCCATCCCGGCGGC 2820
Db 2780 AATGACTGTATGCACAAACAAACGGGCGAGTGTGGGCGAGTGTGCCTTGCCATCCCGGCGGC 2839
QY 2821 CACCGCTGGGCTGGGCTCACACTACACCTTGACACCCAGCAGCCGCAACTGACAGCCG 2880
Db 2840 CACCGCTGGGCTGGGCTCACACTACACCTTGACACCCAGCAGCCGCAACTGACAGCCG 2899
QY 2881 CCCACCACTTCTTGTCTGTTCAGCGAAGAAATCTGCTCATCATCGGATGATTCGGAGAC 2940
Db 2900 CCCACCACTTCTTGTCTGTTCAGCGAAGAAATCTGCTCATCATCGGATGATTCGGAGAC 2959
QY 2941 CAGCACAGCCCGGATCTCATCTCCCTGATCGACTGAGGAAAGTCAAGGACCATCGAC 3000
Db 2960 CAGCACAGCCCGGATCTCATCTCCCTGATCGACTGAGGAAAGTCAAGGACCATCGAC 3019
QY 3001 TATGACCCACATCGACAAAGTTTCATCTACTGGGTGATGGGCGCCAGAAACATCAAGCGAGCC 3060
Db 3020 TATGACCCACATCGACAAAGTTTCATCTACTGGGTGATGGGCGCCAGAAACATCAAGCGAGCC 3079
QY 3061 AAGGACGAGCGGACCCAGCCCTTTTGTGACCTCTCTGAGCCAAAGGCCAAAACCCAGAC 3120
Db 3080 AAGGACGAGCGGACCCAGCCCTTTTGTGACCTCTCTGAGCCAAAGGCCAAAACCCAGAC 3139
QY 3121 AGGACGCCCCACGACCTCAGCATCGACATCTACAGCCGACACACTGTTCTGGACGTGCGAG 3180
Db 3140 AGGACGCCCCACGACCTCAGCATCGACATCTACAGCCGACACACTGTTCTGGACGTGCGAG 3199
QY 3181 GCCACCAATACATCAACAGTCCACAGGCTCAGCGGGGAAGCCATGGGGGTGTGTCGGT 3240
Db 3200 GCCACCAATACATCAACAGTCCACAGGCTCAGCGGGGAAGCCATGGGGGTGTGTCGGT 3259
QY 3241 GGGGACCGCGAACAAGCCAGGGCCATCGTCTGTAACGGCGAGCGAGGCTACTGTTACTTC 3300
Db 3260 GGGGACCGCGAACAAGCCAGGGCCATCGTCTGTAACGGCGAGCGAGGCTACTGTTACTTC 3319
QY 3301 ACCAACATGACGAGCAGCGGACGCAAGATCCAAACGCGAGCCCTTGGACGGCACCGAGCGC 3360
Db 3320 ACCAACATGACGAGCAGCGGACGCAAGATCCAAACGCGAGCCCTTGGACGGCACCGAGCGC 3379
QY 3361 GAGTCTCTTTTACACACCGGCTCATCCGCCCTGTGGCCCTGTGGTGGAGCAACACACTG 3420
Db 3380 GAGTCTCTTTTACACACCGGCTCATCCGCCCTGTGGCCCTGTGGTGGAGCAACACACTG 3439
QY 3421 GGCAGACTTCTGGGTGGAGCGCGACCTGAGCGCATTTGAGAGCTGTGACTGTGACGG 3480
Db 3440 GGCAGACTTCTGGGTGGAGCGCGACCTGAGCGCATTTGAGAGCTGTGACTGTGACGG 3499
QY 3481 GCCAACCGCTGACCTTGGAGGACGCGCAACATCGTGCAGCCCTTGGGCTTGACCATCCTT 3540
Db 3500 GCCAACCGCTGACCTTGGAGGACGCGCAACATCGTGCAGCCCTTGGGCTTGACCATCCTT 3559
QY 3541 GGCAGGATCTTACTGATCGACCGCGCAGCAGAGATGATCGAGCGTGTGGAGAGACC 3600
Db 3560 GGCAGGATCTTACTGATCGACCGCGCAGCAGAGATGATCGAGCGTGTGGAGAGACC 3619
QY 3601 ACCGGGACAAAGCGGACTCGCATCCAGGGCGGTGGCCACCTCATCTGGCATCCATGCA 3660
Db 3620 ACCGGGACAAAGCGGACTCGCATCCAGGGCGGTGGCCACCTCATCTGGCATCCATGCA 3679
QY 3661 GTGGAGGAGTCAAGCTTGGAGGAGTTCACGCCCAACCATGTGCCCGTGACAAATGGTGGC 3720
Db 3680 GTGGAGGAGTCAAGCTTGGAGGAGTTCACGCCCAACCATGTGCCCGTGACAAATGGTGGC 3739
QY 3721 TGCTCCCAACATCTGATTTGCCAAGGGTGAATGGGACACACCGTGTCTATGCCCAGTCCAC 3780
Db 3740 TGCTCCCAACATCTGATTTGCCAAGGGTGAATGGGACACACCGTGTCTATGCCCAGTCCAC 3799
QY 3781 CTGCTGCTCTGCAAGACCTGCTGACCTGTGGAGGCGGCCACCTGCTCCCGGACGAG 3840
Db 3800 CTGCTGCTCTGCAAGACCTGCTGACCTGTGGAGGCGGCCACCTGCTCCCGGACGAG 3859

Db 500 GTGCTCTTCTGGCAGGACCTTGACACGCGGAGGCGCATCGCTTGGACCCCGCTCAACGGG 559
QY 541 TACATGTACTGGACAGACTGGGGTGAAGCGCCCCCGGATGGAGCGGCGAGGATGGC 600
Db 560 TACATGTACTGGACAGACTGGGGTGAAGCGCCCCCGGATGGAGCGGCGAGGATGGC 619
QY 601 AGCACCCGGAAGATCATTTGTGGACTCGGACATTTACTGSCCAATGGAGTGAACCATCGAC 660
Db 620 AGCACCCGGAAGATCATTTGTGGACTCGGACATTTACTGSCCAATGGAGTGAACCATCGAC 679
QY 661 CTGGAGGAGCAAGAGCTCTACTTGGCTGACGCAAGCTCAGCTTCAATCCACGTCGCCAAC 720
Db 680 CTGGAGGAGCAAGAGCTCTACTTGGCTGACGCAAGCTCAGCTTCAATCCACGTCGCCAAC 739
QY 721 CTGGAGGAGCTGTTTCCGCGAGAGGTGGTGGAGGCGAGCTGACGCAACCCCTTGGCCCTG 780
Db 740 CTGGAGGAGCTGTTTCCGCGAGAGGTGGTGGAGGCGAGCTGACGCAACCCCTTGGCCCTG 799
QY 781 ACGCTCTCCGGGACACTCTGTACTTGGACAGACTTGGCAGACCCCGCTCCATCCATGSCCTGC 840
Db 800 ACGCTCTCCGGGACACTCTGTACTTGGACAGACTTGGCAGACCCCGCTCCATCCATGSCCTGC 859
QY 841 AACAGCCGCACTGGGGGAGAGAGAGATCTCTGAGTGCCTCTACTCAACCCATGGAC 900
Db 860 AACAGCCGCACTGGGGGAGAGAGAGATCTCTGAGTGCCTCTACTCAACCCATGGAC 919
QY 901 ATCCAGGTGCTGAGCCAGGAGCGGAGCTTCTTCCACACTCGCTGTGGAGGACAAAT 960
Db 920 ATCCAGGTGCTGAGCCAGGAGCGGAGCTTCTTCCACACTCGCTGTGGAGGAGGACAAAT 979
QY 961 GCGGCTGCTCCCACTGTGCTGCTCCCAAGCAGAGCTTCTTACACATGCGCCTGC 1020
Db 980 GCGGCTGCTCCCACTGTGCTGCTCCCAAGCAGAGCTTCTTACACATGCGCCTGC 1039
QY 1021 CCCAGGCTGTGAGCTGACAGCAACCGGAGAGAGCTGTAAAGGAGGAGGAGCGAGAGGTG 1080
Db 1040 CCCAGGCTGTGAGCTGACAGCAACCGGAGAGAGCTGTAAAGGAGGAGGAGCGAGAGGTG 1099
QY 1081 CTGCTGTGCGCGCGGAGCGGAGCTTACGAGGAGTCTCGCTGGACACGCGGAGCTTAC 1140
Db 1100 CTGCTGTGCGCGCGGAGCGGAGCTTACGAGGAGTCTCGCTGGACACGCGGAGCTTAC 1159
QY 1141 GACATCGTGTGCAAGTGGAGACATCCGCGACGCAATGGCCATCGACTACGACCCGCTA 1200
Db 1160 GACATCGTGTGCAAGTGGAGACATCCGCGACGCAATGGCCATCGACTACGACCCGCTA 1219
QY 1201 GAGGCTATGTCTACTGACAGATGACGAGTGGGGCCATCCGAGGCGGTACTGAGAC 1260
Db 1220 GAGGCTATGTCTACTGACAGATGACGAGTGGGGCCATCCGAGGCGGTACTGAGAC 1279
QY 1261 GGGTCTGGGGCGCAGACGCTGGTCAACACGAGATCAACGACCCCGATGGCATCGCGTC 1320
Db 1280 GGGTCTGGGGCGCAGACGCTGGTCAACACGAGATCAACGACCCCGATGGCATCGCGTC 1339
QY 1321 GACTGGGTGGCCGAAACCTCTACTTGGACCGACACGCGCAGCGGACCGCATCAGGTGACG 1380
Db 1340 GACTGGGTGGCCGAAACCTCTACTTGGACCGACACGCGCAGCGGACCGCATCAGGTGACG 1399
QY 1381 CGCCTCAACGCGACCTCCGCAAGATCTGTGTGTCGAGGAGCTTGGAGGAGCCCGGAGCC 1440
Db 1400 CGCCTCAACGCGACCTCCGCAAGATCTGTGTGTCGAGGAGCTTGGAGGAGCCCGGAGCC 1459
QY 1441 ATCGCACTGCAACCCCGTGTATGGGCTCATGTACTTGGACAGACTGGGAGAGAACCTTAA 1500
Db 1460 ATCGCACTGCAACCCCGTGTATGGGCTCATGTACTTGGACAGACTGGGAGAGAACCTTAA 1519
QY 1501 ATCGAGTGTGCAACTTGGATGGGAGGAGCGGCTGTGTGTCATATGCTCCCTCGG 1560
Db 1520 ATCGAGTGTGCAACTTGGATGGGAGGAGCGGCTGTGTGTCATATGCTCCCTCGG 1579
QY 1561 TGGCCCAACGCGCTGGCCCTGGACCTCGAGGAGGAGGAGCTTACTTGGGAGAGCGCAAG 1620
Db 1580 TGGCCCAACGCGCTGGCCCTGGACCTCGAGGAGGAGGAGCTTACTTGGGAGAGCGCAAG 1639

QY 1621 ACAGCAAGATCGAGGTGATCAATGTTGATGGAGCAAGAGGCGGACCCCTCTCTGGAGGAC 1680
Db 1640 ACAGCAAGATCGAGGTGATCAATGTTGATGGAGCAAGAGGCGGACCCCTCTCTGGAGGAC 1699
QY 1681 AAGCTCCCGCAGCATTTTCGGGTTTCAACGCTGTCTGGGGACTTCTACTCTGAGCTACTG 1740
Db 1700 AAGCTCCCGCAGCATTTTCGGGTTTCAACGCTGTCTGGGGACTTCTACTCTGAGCTACTG 1759
QY 1741 CAGCGCCGACGATCGAGCGGGTGCAAGAGTCAAGGCCAGCCGCGGAGCGTCAATATTGAC 1800
Db 1760 CAGCGCCGACGATCGAGCGGGTGCAAGAGTCAAGGCCAGCCGCGGAGCGTCAATATTGAC 1819
QY 1801 CAGCTGCCCGACCTGATGGGGCTCAAGCTGTGAAATGTGGCAAGGTCTGTGGAAACCAAC 1860
Db 1820 CAGCTGCCCGACCTGATGGGGCTCAAGCTGTGAAATGTGGCAAGGTCTGTGGAAACCAAC 1879
QY 1861 CCGTGTGCGGACAGGAAACGGGGGTGACAGCCACTGTGCTTCTTACACCCCGCAAC 1920
Db 1880 CCGTGTGCGGACAGGAAACGGGGGTGACAGCCACTGTGCTTCTTCAACCCCGCAAC 1939
QY 1921 CCGTGTGCTGCCCTCCCATCGGCTTGGAGCTGTGAGTGAATGAAGACCTGCTGCTGCT 1980
Db 1940 CCGTGTGCTGCCCTCCCATCGGCTTGGAGCTGTGAGTGAATGAAGACCTGCTGCTGCT 1999
QY 1981 GAGGCTTCTTGGTCTTCAACGAGAGCGCGCATCCACAGGATCTCCCTCGAGACCAAT 2040
Db 2000 GAGGCTTCTTGGTCTTCAACGAGAGCGCGCATCCACAGGATCTCCCTCGAGACCAAT 2059
QY 2041 AACCAACGAGCTGGGCTATCCGCTCAACGAGGCTCAAGAGGCTCAGCCCTGAGCTTTGAT 2100
Db 2060 AACCAACGAGCTGGGCTATCCGCTCAACGAGGCTCAAGAGGCTCAGCCCTGAGCTTTGAT 2119
QY 2101 GTGTCCAAACCAACATCTACTGACAGAGCTCAGCTTGAAGACCATCAGCGCGGCTTC 2160
Db 2120 GTGTCCAAACCAACATCTACTGACAGAGCTCAGCTTGAAGACCATCAGCGCGGCTTC 2179
QY 2161 ATGAAACGAGCTCGGTGGAGCAGCTGTGGAGTTTGGCTTGTACTACCCCGAGGCGATG 2220
Db 2180 ATGAAACGAGCTCGGTGGAGCAGCTGTGGAGTTTGGCTTGTACTACCCCGAGGCGATG 2239
QY 2221 GCCCTTGACTGGATGGGCAAGAACCTCTACTGGGCGGACACTGGGACCAACAGAAATCGAA 2280
Db 2240 GCCCTTGACTGGATGGGCAAGAACCTCTACTGGGCGGACACTGGGACCAACAGAAATCGAA 2299
QY 2281 GTGCGCGGCTGAGCGGCGAGTTCCGCGAGTCTCGTGTGGAGGAGCTTGGACCAACCCG 2340
Db 2300 GTGCGCGGCTGAGCGGCGAGTTCCGCGAGTCTCGTGTGGAGGAGCTTGGACCAACCCG 2359
QY 2341 AGGTGCTGCGCTTGGATCCCAACAGGCTTACATCTACTTGGACCGAGTGGGCGGCAAG 2400
Db 2360 AGGTGCTGCGCTTGGATCCCAACAGGCTTACATCTACTTGGACCGAGTGGGCGGCAAG 2419
QY 2401 CCGAGGATCGTGGGCGCTTTCATGGACGGGACCAACTGATGACGCTGGTGGACCAAGTG 2460
Db 2420 CCGAGGATCGTGGGCGCTTTCATGGACGGGACCAACTGATGACGCTGGTGGACCAAGTG 2479
QY 2461 GCGCGGCGCAACGACCTTCAACATTTGACTACGCTGACGAGCGCTTCTACTTGGACCGACTG 2520
Db 2480 GCGCGGCGCAACGACCTTCAACATTTGACTACGCTGACGAGCGCTTCTACTTGGACCGACTG 2539
QY 2521 GACACCAACATGATCGAGTCTGCAACATGCTGGGTGAGGCGGGTCTGTATTGGCGGAC 2580
Db 2540 GACACCAACATGATCGAGTCTGCAACATGCTGGGTGAGGCGGGTCTGTATTGGCGGAC 2599
QY 2581 GATCTCCCGCAGCCGCTTGTGAGCGAGTACAGCGATTATATCTACTTGGACAGACTGG 2640
Db 2600 GATCTCCCGCAGCCGCTTGTGAGCGAGTACAGCGATTATATCTACTTGGACAGACTGG 2659
QY 2641 AATCTGACAGCTGAGCGGCGGCAAGACTAGCGGCGGAGCCGACCCCTCATCCAG 2700
Db 2660 AATCTGACAGCTTGGAGCGGCGGCAAGACTAGCGGCGGAGCCGACCCCTCATCCAG 2719

Qy	2701	GGCCACCTGGACTTCGTGATGGACAATCCTGTGTGTTTCACTCTCTCCCGCAGGATGGCCTC	2761
Db	2720	GGCCACCTGGACTTCGTGATGGACAATCCTGTGTGTTTCCACTCTCTCCCGCAGGATGGCCTC	2779
Qy	2761	AATGACTGTATGCAACAACCGGCAGTGTGGCAGCTGTGTGCTTGGCATNTCCCGCGCGGC	2820
Db	2780	AATGACTGTATGCAACAACCGGCAGTGTGGCAGCTGTGTGCTTGGCATNTCCCGCGCGGC	2839
Qy	2821	CACCGCTGCGGCTCGGCTTCACTACAACCTTGGACCCGACAGCCGCAACTGTGGATGATCCGGAGCAC	2880
Db	2840	CACCGCTGCGGCTCGGCTTCACTACAACCTTGGACCCGACAGCCGCAACTGTGGATGATCCGGAGCAC	2899
Qy	2881	CCCACACCTTCTTGTCTGTTTCAGCCAGAAATCTGCACTCAGTGGATGATCCGGAGCAC	2940
Db	2900	CCCACACCTTCTTGTCTGTTTCAGCCAGAAATCTGCACTCAGTGGATGATCCGGAGCAC	2959
Qy	2941	CAGCAGCCGCGATCTCATCTCTGCCCTGTGATGCACTGAGAAACGTCAAAGCCATCGAC	3000
Db	2960	CAGCAGCCGCGATCTCATCTCTGCCCTGTGATGCACTGAGAAACGTCAAAGCCATCGAC	3019
Qy	3001	TATGACCCACTGGCAAAAGTTTCACTACTGTGGTGGATGGCGCCAGAAACATCAAGCGAGCC	3060
Db	3020	TATGACCCACTGGCAAAAGTTTCACTACTGTGGTGGATGGCGCCAGAAACATCAAGCGAGCC	3079
Qy	3061	AAGGACAGCGGACCCAGCCCTTGTGTTTGTGACCTCTCTGAGCCAGAGCCAAAACCCAGAC	3120
Db	3080	AAGGACAGCGGACCCAGCCCTTGTGTTTGTGACCTCTCTGAGCCAGAGCCAAAACCCAGAC	3139
Qy	3121	AGGAGAGCCACAGACTCAGATCGACATCTACAGCCGAGACATGTGTTCTGAGAGTGGAG	3180
Db	3140	AGGAGAGCCACAGACTCAGATCGACATCTACAGCCGAGACATGTGTTCTGAGAGTGGAG	3199
Qy	3181	GCCACCAATACCATCAACGCTCCAGGCTGAGCGGGGAAGCCATGGGGGTGGTGTCTGCTCGT	3240
Db	3200	GCCACCAATACCATCAACGCTCCAGGCTGAGCGGGGAAGCCATGGGGGTGGTGTCTGCTCGT	3259
Qy	3241	GGGGAACCGGACAAGCCACGAGCCATCGTGTGTCAAACGCGAGCGAGAGGTACTGTACTTC	3300
Db	3260	GGGGAACCGGACAAGCCACGAGCCATCGTGTGTCAAACGCGAGCGAGAGGTACTGTACTTC	3319
Qy	3301	ACCAAATGACAGACCGGGCAGCCAAAGATCGAACGGCAGCCCTTGGACGACCGAGCGC	3360
Db	3320	ACCAAATGACAGACCGGGCAGCCAAAGATCGAACGGCAGCCCTTGGACGACCGAGCGC	3379
Qy	3361	GAGTCTCTTTACCAACCGGCTCATCCGCTGTGGCCCTGTGTGGTGGAGCAACACTG	3420
Db	3380	GAGTCTCTTTACCAACCGGCTCATCCGCTGTGGCCCTGTGTGGTGGAGCAACACTG	3439
Qy	3421	GGCAAGCTGTTTGGGTGGACGGGACCTGAAGCGCAATGAGAGCTGTGACTGTGAGGG	3480
Db	3440	GGCAAGCTGTTTGGGTGGACGGGACCTGAAGCGCAATGAGAGCTGTGACTGTGAGGG	3499
Qy	3481	GCCAAACGCTGACCTGGAGGAGCCAAATCGTGGAGCCTCTGGGCTGTGACCATCTT	3540
Db	3500	GCCAAACGCTGACCTGGAGGAGCTTCTGAGCCACCCATCGTGGAGCCTCTGGGCTGTGACCATCTT	3559
Qy	3541	GGCAAGCATCTCTACTGGATCGACCGGACGAGCAGATGATCGAGCGTGTGGAGAGAGCC	3600
Db	3560	GGCAAGCATCTCTACTGGATCGACCGGACGAGCAGATGATCGAGCGTGTGGAGAGAGCC	3619
Qy	3601	ACCGGGACAAGCGGACTCGCATCCAGGGCCGTGTGCCCACTCACTGGCATCCATGCA	3660
Db	3620	ACCGGGACAAGCGGACTCGCATCCAGGGCCGTGTGCCCACTCACTGGCATCCATGCA	3679
Qy	3661	GTGGAGGAAGTCAAGCTGGAGGAGTTCTCAGCCACCCATGTGCCCGTGTGAATGTGTGC	3720
Db	3680	GTGGAGGAAGTCAAGCTGGAGGAGTTCTCAGCCACCCATGTGCCCGTGTGAATGTGTGC	3739
Qy	3721	TGCTCCCAACATCTGTATTGCAAGGGTGTGGGACACACAGGTGTCTCATGCCAGATCCAC	3780
Db	3740	TGCTCCCAACATCTGTATTGCAAGGGTGTGGGACACACAGGTGTCTCATGCCAGATCCAC	3799
Qy	3781	CTGTGTCTCTGCAAGACCTGTGACTGTGGAGAGCGCGCCACCTGCTCTCCCGGACCCAG	3840

[illegible]

QY 424 ACAGACTCAGAGACCAACCGCATCGAGTGGCCAACTCAATGACACATCCCGAAGGTG 483
DB |||||
QY 421 ACAGACTCAGAGACCAACCGCATCGAGTGGCCAACTCAATGACACATCCCGAAGGTG 480
DB |||||
QY 484 CTCCTTGGCAGGACCTTGAACAGCTAGAGCCATCGCTTTGGACCCCGCTCACGGGTAC 543
DB |||||
QY 481 CTCCTTGGCAGGACCTTGAACAGCTAGAGCCATCGCTTTGGACCCCGCTCACGGGTAC 540
DB |||||
QY 544 ATGTACTGGACAGACTGGGTGAGACCGCCCGGATTGAGCGGGCAGGGATGGTGCAGC 603
DB |||||
QY 541 ATGTACTGGACAGACTGGGTGAGACCGCCCGGATTGAGCGGGCAGGGATGGTGCAGC 600
DB |||||
QY 604 ACCCGAAGATCATTTGTGGACTCGGACATTTACTTGGCCCAATGACTGACCATCGACTG 663
DB |||||
QY 601 ACCCGAAGATCATTTGTGGACTCGGACATTTACTTGGCCCAATGACTGACCATCGACTG 660
DB |||||
QY 664 GAGGACGAGAGCTCTACTGGGCTGAGCCCAAGCTCAGCTTTCATCCACCGTGCCAACTG 723
DB |||||
QY 661 GAGGACGAGAGCTCTACTGGGCTGAGCCCAAGCTCAGCTTTCATCCACCGTGCCAACTG 720
DB |||||
QY 724 GACGGCTGTTCCGGCAGAGGTGGTGGAGGCGACCTTGACGACCCCTTGCCTGACG 783
DB |||||
QY 721 GACGGCTGTTCCGGCAGAGGTGGTGGAGGCGACCTTGACGACCCCTTGCCTGACG 780
DB |||||
QY 784 CTCTCCGGGACACTCTGTACTGGACAGACTGGCAGACCCGCTCCATCCATGCTGCAAC 843
DB |||||
QY 781 CTCTCCGGGACACTCTGTACTGGACAGACTGGCAGACCCGCTCCATCCATGCTGCAAC 840
DB |||||
QY 844 AAGCGCACTGGGGGAGAGAGAGATCTGTAGTGCCCTCTACTCACCCATGGACATC 903
DB |||||
QY 841 AAGCGCACTGGGGGAGAGAGAGATCTGTAGTGCCCTCTACTCACCCATGGACATC 900
DB |||||
QY 904 CAGTGCTGACGACGAGGCGCAGCTTCTTTCACACTCGCTGTGAGGAGCAATGGC 963
DB |||||
QY 901 CAGTGCTGACGACGAGGCGCAGCTTCTTTCACACTCGCTGTGAGGAGCAATGGC 960
DB |||||
QY 964 GGCTGCTCCCACTGTGCTCTGTCCCAAGCAGCTTCTTACATGCGCTGCCCC 1023
DB |||||
QY 961 GGCTGCTCCCACTGTGCTCTGTCCCAAGCAGCTTCTTACATGCGCTGCCCC 1020
DB |||||
QY 1024 ACGGTGTGACGTGACGAGCAACCGCAGGACGTGTAAAGCAGGACCGGAGGAGTGTG 1083
DB |||||
QY 1021 ACGGTGTGACGTGACGAGCAACCGCAGGACGTGTAAAGCAGGACCGGAGGAGTGTG 1080
DB |||||
QY 1084 CTGCTGGCCCGGCGGACGAGACTAGGAGGATCTGCTGGACACGCGGAGTTCACCGAC 1143
DB |||||
QY 1081 CTGCTGGCCCGGCGGACGAGACTAGGAGGATCTGCTGGACACGCGGAGTTCACCGAC 1140
DB |||||
QY 1144 ATCGTGTGACAGTGGACGACATCCGCGACGCCATTCATCGACTAGACCCGCTAGAG 1203
DB |||||
QY 1141 ATCGTGTGACAGTGGACGACATCCGCGACGCCATTCATCGACTAGACCCGCTAGAG 1200
DB |||||
QY 1204 GGCTATGTCTACTGGACAGATGACGAGTGGCGGCCATTCGCGAGGCGGTACTGGACGGG 1263
DB |||||
QY 1201 GGCTATGTCTACTGGACAGATGACGAGTGGCGGCCATTCGCGAGGCGGTACTGGACGGG 1260
DB |||||
QY 1264 TCTGGGGCGAGACGCTGGTCAACACCGAGATCAACGACCCCGATGGCATGCGGTCGAC 1323
DB |||||
QY 1261 TCTGGGGCGAGACGCTGGTCAACACCGAGATCAACGACCCCGATGGCATGCGGTCGAC 1320
DB |||||
QY 1324 TGGGTGGCCGAAACCTTCTACTGGACGACACGCGGCGACGCGCATCGAGGTGACGCGC 1383
DB |||||
QY 1321 TGGGTGGCCGAAACCTTCTACTGGACGACACGCGGCGACGCGCATCGAGGTGACGCGC 1380
DB |||||
QY 1384 CTCACGCGCACTCCCGCAAGATCTGTGTGGAGGACCTTGGACGAGCCCGCGGACCATC 1443
DB |||||
QY 1381 CTCACGCGCACTCCCGCAAGATCTGTGTGGAGGACCTTGGACGAGCCCGCGGACCATC 1440
DB |||||
QY 1444 GCACTGCAACCCGTGATGGGCTCATGTACTGGACAGACTGGGAGAGAACCTTAAATC 1503
DB |||||
QY 1441 GCACTGCAACCCGTGATGGGCTCATGTACTGGACAGACTGGGAGAGAACCTTAAATC 1500
DB |||||

QY 1504 GAGTGTGCCAACTTGGATGGGACGCGCTGTGTGGTCAATGCTCCTCGGGTGG 1563
DB |||||
QY 1501 GAGTGTGCCAACTTGGATGGGACGCGCTGTGTGTCAATGCTCCTCGGGTGG 1560
DB |||||
QY 1564 CCCAAACCGGCTGGCCCTGGACCTGACGAGAGGGAGCTCTACTGGGAGAGACCCAAAGACA 1623
DB |||||
QY 1561 CCCAAACCGGCTGGCCCTGGACCTGACGAGAGGGAGCTCTACTGGGAGAGACCCAAAGACA 1620
DB |||||
QY 1624 GACAAGATCAGAGGTGATCAATGTTGATGGGACGAGAGGGCGGACCTCCTCGAGGACAAAG 1683
DB |||||
QY 1621 GACAAGATCAGAGGTGATCAATGTTGATGGGACGAGAGGGCGGACCTCCTCGAGGACAAAG 1680
DB |||||
QY 1684 CTCGCCGACATTTTCGGGTTTTCAGCTGTGGGGGACTTTCATCTACTGACTGACTGGCAG 1743
DB |||||
QY 1681 CTCGCCGACATTTTCGGGTTTTCAGCTGTGGGGGACTTTCATCTACTGACTGACTGGCAG 1740
DB |||||
QY 1744 CGCCGACGATCGAGCGGGTGCAAAAGGTCAAGGCCAGCCCGGGACGCTCATCATTTGACCAAG 1803
DB |||||
QY 1741 CGCCGACGATCGAGCGGGTGCAAAAGGTCAAGGCCAGCCCGGGACGCTCATCATTTGACCAAG 1800
DB |||||
QY 1804 CTGCCCGACCTGTAGTGGGGCTCAAAAGCTGTGAATGTGGCCAAAGGTGCTCGGAAACCAACCGG 1863
DB |||||
QY 1801 CTGCCCGACCTGTAGTGGGGCTCAAAAGCTGTGAATGTGGCCAAAGGTGCTCGGAAACCAACCGG 1860
DB |||||
QY 1864 TGTGGCGACAGGAAACGGGGGGTGAGCCACTGTGTCTTTCACACCCCGCAACCCGG 1923
DB |||||
QY 1861 TGTGGCGACAGGAAACGGGGGGTGAGCCACTGTGTCTTTCACACCCCGCAACCCGG 1920
DB |||||
QY 1924 TGTGGCTGCCCCATTCGGCCCTGGAGCTGTGTGATGACATGAAGACTGTGATCGTGGCTGAG 1983
DB |||||
QY 1921 TGTGGCTGCCCCATTCGGCCCTGGAGCTGTGTGATGACATGAAGACTGTGATCGTGGCTGAG 1980
DB |||||
QY 1984 GCCTTCTTGTCTTTCACACGAGGAGCGCCATTCACAGAGTCTCCCTCGAGACCAATTAAC 2043
DB |||||
QY 1981 GCCTTCTTGTCTTTCACACGAGGAGCGCCATTCACAGAGTCTCCCTCGAGACCAATTAAC 2040
DB |||||
QY 2044 AACGACGTGGCCATTCGCCCTCACGGGGCTCAAGGAGCGCTCAGGCCCTGGACTTTGATGTG 2103
DB |||||
QY 2041 AACGACGTGGCCATTCGCCCTCACGGGGCTCAAGGAGCGCTCAGGCCCTGGACTTTGATGTG 2100
DB |||||
QY 2104 TCCAAACCAACATCTACTGAGACAGAGCTCAGCCCTGAAGACCATCAGCGCGGCTTCATG 2163
DB |||||
QY 2101 TCCAAACCAACATCTACTGAGACAGAGCTCAGCCCTGAAGACCATCAGCGCGGCTTCATG 2160
DB |||||
QY 2164 AACGGGAGCTCGGTGGAGCAGCTGGTGGAGTTTGGCCCTTGACTACCCCGAGGGATGGCC 2223
DB |||||
QY 2161 AACGGGAGCTCGGTGGAGCAGCTGGTGGAGTTTGGCCCTTGACTACCCCGAGGGATGGCC 2220
DB |||||
QY 2224 GTTGACTGGATGGGCAAGAACCTCTACTGGGCGGACACTGGGACCAACAGAAATCGAAGTG 2283
DB |||||
QY 2221 GTTGACTGGATGGGCAAGAACCTCTACTGGGCGGACACTGGGACCAACAGAAATCGAAGTG 2280
DB |||||
QY 2284 GCGCGCTGAGACGGGCAAGTTCCGGCAAGTCTCTGTGTGGAGGGAATTGGGAACAACCGGAGG 2343
DB |||||
QY 2281 GCGCGCTGAGACGGGCAAGTTCCGGCAAGTCTCTGTGTGGAGGGAATTGGGAACAACCGGAGG 2340
DB |||||
QY 2344 TCGTGGCCCTGGATCCCAACGAGGCTTACATCTACTGGACCGGAGTGGGGCGGCAAGCGG 2403
DB |||||
QY 2341 TCGTGGCCCTGGATCCCAACGAGGCTTACATCTACTGGACCGGAGTGGGGCGGCAAGCGG 2400
DB |||||
QY 2404 AGGATCTGGGGGCTTTCATGGACGGGACCAATTCGATGACGCTGGTGGACCAAGGTGGGC 2463
DB |||||
QY 2401 AGGATCTGGGGGCTTTCATGGACGGGACCAATTCGATGACGCTGGTGGACCAAGGTGGGC 2460
DB |||||
QY 2464 CGGGCCAAACGACTTCAACATTTGACTACGCTGACAGCGCTCTTACTGGAACCGACTGGAC 2523
DB |||||
QY 2461 CGGGCCAAACGACTTCAACATTTGACTACGCTGACAGCGCTCTTACTGGAACCGACTGGAC 2520
DB |||||
QY 2524 ACCAAGATGATCGAGTGTGCTCCACATCTGGGTGAGGAGCGGCTGTGATTTGCCGACAT 2583
DB |||||
QY 2521 ACCAAGATGATCGAGTGTGCTCCACATCTGGGTGAGGAGCGGCTGTGATTTGCCGACAT 2580
DB |||||
QY 2584 CTCGCCGACCGGCTTTCGGTCTGACGAGTACAGCGATTATATCTACTGGACAGACTGGAAAT 2643
DB |||||

Qy	1501	ATCAGTGTGCAACTTGGATGGGACGAGCGGCGTGTGCTGGTCAATGCTCCTCGGG	1560	Db	2600	GAITCCCGCACCCGTTTCGGTCTGACGCACTATATCTACTTGACAGACTGG	2659
Db	1520	ATCAGTGTGCAACTTGGATGGGACGAGCGGCGTGTGCTGGTCAATGCTCCTCGGG	1579	Qy	2641	AATCTGACAGCAATTTAGCGGGCGGCAAGACATAGCGGCGGGAACCGACCTCATCCAG	2700
Qy	1561	TGGCCCAACGGCCTGGCCCTGGACCTGACAGAGGGGAAGCTCTACTGGGGAGAGCCCAAG	1620	Db	2660	AATCTGACAGCAATTTAGCGGGCGGCAAGACATAGCGGCGGGAACCGACCTCATCCAG	2719
Db	1580	TGGCCCAACGGCCTGGCCCTGGACCTGACAGAGGGGAAGCTCTACTGGGGAGAGCCCAAG	1639	Qy	2701	GGCCACCTGACATTCGTGATGGACATCTCGTGTTCACCTCTCCCGCAGGATGGGCTC	2760
Qy	1621	ACAGACAAGATCGAGGTGATCAATGTTGATGGGACGAAGAGCGGGAACCTCTCTGGAGAC	1680	Db	2720	GGCCACCTGACATTCGTGATGGACATCTCGTGTTCACCTCTCCCGCAGGATGGGCTC	2779
Db	1640	ACAGACAAGATCGAGGTGATCAATGTTGATGGGACGAAGAGCGGGAACCTCTCTGGAGAC	1699	Qy	2761	AATGACTGTATGCACAAACGGGCACTGTGGGCACTGTGCTTGGCATCCCCGGGGC	2820
Qy	1681	AAGCTCCCGCACATTTTCGGGTTCAAGCTGTGGGGACCTTCACTACTGACATGACTGG	1740	Db	2780	AATGACTGTATGCACAAACGGGCACTGTGGGCACTGTGCTTGGCATCCCCGGGGC	2839
Db	1700	AAGCTCCCGCACATTTTCGGGTTCAAGCTGTGGGGACCTTCACTACTGACATGACTGG	1759	Qy	2821	CACCGCTGGGGCTGCGCTCACACTACCTTGAGACCCGAGCAGCGAACTGACGCGC	2880
Qy	1741	CAGCGCGCAGCATCGAGCGGGTGCAAAAGGTCAAGCGCAGCGGGAGCTCATCATTTGAC	1800	Db	2840	CACCGCTGGGGCTGCGCTCACACTACCTTGAGACCCGAGCAGCGAACTGACGCGC	2899
Db	1760	CAGCGCGCAGCATCGAGCGGGTGCAAAAGGTCAAGCGCAGCGGGAGCTCATCATTTGAC	1819	Qy	2881	CCACACCTCTTGTGTTTCAGCCAGAAATCTGCCATCAGTCGAGATGATCCCGGACGAC	2940
Qy	1801	CAGTGTCCCGACCTGATGGGGCTCAAGCTGTGAATGTGGCCAAAGGTGCTGCGAAACCAAC	1860	Db	2900	CCACACCTCTTGTGTTTCAGCCAGAAATCTGCCATCAGTCGAGATGATCCCGGACGAC	2959
Db	1820	CAGTGTCCCGACCTGATGGGGCTCAAGCTGTGAATGTGGCCAAAGGTGCTGCGAAACCAAC	1879	Qy	2941	CAGCACAGCCCGGATCTCATCTCTGCCCTGCACTGAGACTGAGGAACGTCAAAGCCATCGAC	3000
Qy	1861	CCGTGTGGGACAGAAACGGGGGTGACAGCCACCTGTGCTTTCACACCCACGCAACC	1920	Db	2960	CAGCACAGCCCGGATCTCATCTCTGCCCTGCACTGAGACTGAGGAACGTCAAAGCCATCGAC	3019
Db	1880	CCGTGTGGGACAGAAACGGGGGTGACAGCCACCTGTGCTTTCACACCCACGCAACC	1939	Qy	3001	TATGACCCACTGGAACAAGTTTCATCTACTGGTGGATGGGCGCCAGAACTCAAGGAGCC	3060
Qy	1921	CGTGTGCTGCCCATCGGCTCGAGCTGTGAGTGACATGAAGACCTGATCGTGCCT	1980	Db	3020	TATGACCCACTGGAACAAGTTTCATCTACTGGTGGATGGGCGCCAGAACTCAAGGAGCC	3079
Db	1940	CGTGTGCTGCCCATCGGCTCGAGCTGTGAGTGACATGAAGACCTGATCGTGCCT	1999	Qy	3061	AAGGACGACGGGACCCAGCCCTTTGTTTGTGACCTCTCTGAGCCAAAGGCAAAACCCAGAC	3120
Qy	1981	GAGGCTTCTTGGTCTTCAACAGAGAGCCGCGATTCACAGAGATCTCCCTCGAGACCAAT	2040	Db	3080	AAGGACGACGGGACCCAGCCCTTTGTTTGTGACCTCTCTGAGCCAAAGGCAAAACCCAGAC	3139
Db	2000	GAGGCTTCTTGGTCTTCAACAGAGAGCCGCGATTCACAGAGATCTCCCTCGAGACCAAT	2059	Qy	3121	AGGACGACCCCAACGACCTCAGCATCGACATCTACAGCCGGAACACTGTTCTGGAACGTGGAG	3180
Qy	2041	AACAGAGCTGGCCATCCCGCTCAGCGGCTCAAGAGGCTCAGCCCTCGACTTTGAT	2100	Db	3140	AGGACGACCCCAACGACCTCAGCATCGACATCTACAGCCGGAACACTGTTCTGGAACGTGGAG	3199
Db	2060	AACAGAGCTGGCCATCCCGCTCAGCGGCTCAAGAGGCTCAGCCCTCGACTTTGAT	2119	Qy	3181	GGCAACCAATACCAATCAACGCTCAGAGCTGAGCGGGGAGCCATGGGGGTGCTGCGT	3240
Qy	2101	GTGTCCAAACACCACTACTTGGACAGACGCTCAGCTTGAAGACCATCAGCGCGCCTTC	2160	Db	3200	GGCAACCAATACCAATCAACGCTCAGAGCTGAGCGGGGAGCCATGGGGGTGCTGCGT	3259
Db	2120	GTGTCCAAACACCACTACTTGGACAGACGCTCAGCTTGAAGACCATCAGCGCGCCTTC	2179	Qy	3241	GGGACCGGCAAGCCCGAGGCTCATCGTCTCAACGGAGCGAGGCTACTGCTACTTC	3300
Qy	2161	ATGAACGGGAGCTCGGTGGAGCACTGTGTGAGTTGGCTTGTACTACCCGAGGGCATG	2220	Db	3260	GGGACCGGCAAGCCCGAGGCTCATCGTCTCAACGGAGCGAGGCTACTGCTACTTC	3319
Db	2180	ATGAACGGGAGCTCGGTGGAGCACTGTGTGAGTTGGCTTGTACTACCCGAGGGCATG	2239	Qy	3301	ACCAACATGACAGGACCGGGGACGCAAGATCGAAACGGCAGCCCTGGAACCGAGCGC	3360
Qy	2221	GCCGTTGACTGGATGGGCAAGAACCTCTACTGGGCGGACACTGGGACCAACAGAAATCGAA	2280	Db	3320	ACCAACATGACAGGACCGGGGACGCAAGATCGAAACGGCAGCCCTGGAACCGAGCGC	3379
Db	2240	GCCGTTGACTGGATGGGCAAGAACCTCTACTGGGCGGACACTGGGACCAACAGAAATCGAA	2299	Qy	3361	GAGTCTCTTTCACCAACGGGCTCATCGCCCTGTGGCCCTGTGGTGGACAAACACTG	3420
Qy	2281	GTGGCGGGCTGGACGGGAGTTCCGGCAAGTCTCGTGTGGAGGGAATTGGACAACCCG	2340	Db	3380	GAGTCTCTTTCACCAACGGGCTCATCGCCCTGTGGCCCTGTGGTGGACAAACACTG	3439
Db	2300	GTGGCGGGCTGGACGGGAGTTCCGGCAAGTCTCGTGTGGAGGGAATTGGACAACCCG	2359	Qy	3421	GGCAAGCTGTTCTGGGTGGACCGGGAACCTGAAGCGCAATTGAGAGCTGTGACCTGTGAGG	3480
Qy	2341	AGGTCTGGCCCTGGATCCCAACAGGGCTACATCTACTGGACCGAGTGGGGCGGCAAG	2400	Db	3440	GGCAAGCTGTTCTGGGTGGACCGGGAACCTGAAGCGCAATTGAGAGCTGTGACCTGTGAGG	3499
Db	2360	AGGTCTGGCCCTGGATCCCAACAGGGCTACATCTACTGGACCGAGTGGGGCGGCAAG	2419	Qy	3481	GGCAACCGCCTGACCTCTGGAGGAGCGCAACATCGTGCAGCCCTGAGCCCTGACCACTCCTT	3540
Qy	2401	CCGAGGATCTGTGGGGCTTCAATGACCGGACCAACTGACGCTGCTGAGCAAGGTG	2460	Db	3500	GGCAACCGCCTGACCTCTGGAGGAGCGCAACATCGTGCAGCCCTGAGCCCTGAGCCATCCTT	3559
Db	2420	CCGAGGATCTGTGGGGCTTCAATGACCGGACCAACTGACGCTGCTGAGCAAGGTG	2479	Qy	3541	GGCAAGCATCTTACTGGATCGCCAGCAGCAGATGATCGAGCGTGTGGAGAGGACC	3600
Qy	2461	GGCGGGCAACGACCTCAACATGACCTACGCTGACAGCGCTTACTTGAACCGACCTG	2520	Db	3560	GGCAAGCATCTTACTGGATCGCCAGCAGCAGATGATCGAGCGTGTGGAGAGGACC	3619
Db	2480	GGCGGGCAACGACCTCAACATGACCTACGCTGACAGCGCTTACTTGAACCGACCTG	2539	Qy	3601	ACCGGGGCAAGAGCGACTCGCATCCAGGGCGGTGTGCGCCACCTCACTGGCATTCATGCA	3660
Qy	2521	GACACCAACATGATCGAGTCTGCAACATGCTGGGTGAGGAGCGGTGTGATTCGCGAC	2580	Db	3620	ACCGGGGCAAGAGCGACTCGCATCCAGGGCGGTGTGCGCCACCTCACTGGCATTCATGCA	3679
Db	2540	GACACCAACATGATCGAGTCTGCAACATGCTGGGTGAGGAGCGGTGTGATTCGCGAC	2599	Qy	3661	GTGGAGGAGTCAAGCTGGAGGAGTTCTCAGCCCAACCATGTGCGCGTGACAACTGTGCG	3720
Qy	2581	GATCTCCCGCACCCGTTTCGGTCTGACGAGTACAGCGATTATATCTACTGGACAGACTGG	2640				

Qy	361	ATCTCCGGCTGTGGTCTCTCTCCCGACGGCTCTCCCTCGCATGGGTGGGCAAGAAGCTGTAC	420
Db	380	ATCTCCGGCTGTGGTCTCTCTCCCGACGGCTCTCCCTCGCATGGGTGGGCAAGAAGCTGTAC	439
Qy	421	TGGACGGACTCAGAGACCAACCGCATCGAGGTGGCCAACTCAATGGCACATCCCGGAG	480
Db	440	TGGACGGACTCAGAGACCAACCGCATCGAGGTGGCCAACTCAATGGCACATCCCGGAG	499
Qy	481	GTGCTCTTCTGGCAGGACCTTTGACAGCTAGGGCCATCGCTTTGGACCCCGCTCACGGG	540
Db	500	GTGCTCTTCTGGCAGGACCTTTGACAGCCGAGGGCCATCGCTTTGGACCCCGCTCACGGG	559
Qy	541	TACATGTACTGGA CAGACTGGGTGGAGACGCCGCCGATTTGAGCGGGCAGGATGGATGGC	600
Db	560	TACATGTACTGGA CAGACTGGGTTTGAGACGCCGCCGATTTGAGCGGGCAGGATGGATGGC	619
Qy	601	AGCACCCGGGAAGATCATTTGTGGACTCGGACATTTTACTGGGCCAATGGACTGACCATCGAC	660
Db	620	AGCACCCGGGAAGATCATTTGTGGACTCGGACATTTTACTGGGCCAATGGACTGACCATCGAC	679
Qy	661	CTGAGAGACGAAGCTCTACTGGGCTGA CGCCAAAGCTCAGCTTTCATCCACGFGCCAAAC	720
Db	680	CTGAGAGACGAAGCTCTACTGGGCTGAGCCAAAGCTCAGCTTTCATCCACGFGCCAAAC	739
Qy	721	CTGACCGGCTGTTTCCGGCAGAAGTGTGTGGAGGGCAGCTGACGCAACCCCTTCGCCCTG	780
Db	740	CTGACCGGCTGTTTCCGGCAGAAGTGTGTGGAGGGCAGCTGACGCAACCCCTTCGCCCTG	799
Qy	781	ACGCTCTCCGGGGA CACTCTGTACTGGA CAGACTGGCAGACCCGCTCCATCTCATGCTCTG	840
Db	800	ACGCTCTCCGGGGA CACTCTGTACTGGA CAGACTGGCAGACCCGCTCCATCTCATGCTCTG	859
Qy	841	AACAAGCGCACTGGGGGGAAGAGAAAGAGATCCTGTAGTGCCTCTACTCACCCATGGAC	900
Db	860	AACAAGCGCACTGGGGGGAAGAGAAAGAGATCCTGTAGTGCCTCTACTCACCCATGGAC	919
Qy	901	ATCCAGGTGCTGAGCCAGGACGGCAGCCTTTTCTTCCACACTCGCTGTGAGGAGACAAT	960
Db	920	ATCCAGGTGCTGAGCCAGGACGGCAGCCTTTTCTTCCACACTCGCTGTGAGGAGACAAT	979
Qy	961	GGGGGCTCTCCACCTGTGGCTGCTGTCTCCCAAGCAGAGCTTTCATACATGGGCTCTG	1020
Db	980	GGGGGCTCTCCACCTGTGGCTGCTGTCTCCCAAGCAGAGCTTTCATACATGGGCTCTG	1039
Qy	1021	CCACGGGTGTGACGTGTCAGGA CAAACGGCAGGACGTGTAGGCAAGAGCCGAGAGGTG	1080
Db	1040	CCACGGGTGTGACGTGTCAGGA CAAACGGCAGGACGTGTAGGCAAGAGCCGAGAGGTG	1099
Qy	1081	CTGCTGTGGCCCGGCGGA CGGACCTACGAGAGATCTCGCTGGACA CGCCGGA CTTCAAC	1140
Db	1100	CTGCTGTGGCCCGGCGGA CGGACCTACGAGAGATCTCGCTGGACA CGCCGGA CTTCAAC	1159
Qy	1141	GACATCTGTGTGAGGTGGACGACATCCGGCAGCCCAATGCGATCGACTACGACCCGCTA	1200
Db	1160	GACATCTGTGTGAGGTGGACGACATCCGGCAGCCCAATGCGATCGACTACGACCCGCTA	1219
Qy	1201	GAGGGCTATGTCTACTGGA CAGATGACGAGGTGCGGCCATTCGCGAGGGCGTACTCTGGAC	1260
Db	1220	GAGGGCTATGTCTACTGGA CAGATGACGAGGTGCGGCCATTCGCGAGGGCGTACTCTGGAC	1279
Qy	1261	GGGTCTGGGGCGAGACGCTGTCTAACACCGAGATCAACGACCCCGATGGCATCGCGGTC	1320
Db	1280	GGGTCTGGGGCGAGACGCTGTCTAACACCGAGATCAACGACCCCGATGGCATCGCGGTC	1339
Qy	1321	GACTGGGTGGCCGAAACCTCTACTGGA CCGACA CGGGCGACGGACCGCATCTGAGGTGACG	1380
Db	1340	GACTGGGTGGCCGAAACCTCTACTGGA CCGACA CGGGCGACGGACCGCATCTGAGGTGACG	1399
Qy	1381	CGCTCAACGGCACCTCCGCAAGATCCTGGTGTGAGAGACTGTGACGAGCCCCGAGCC	1440
Db	1400	CGCTCAACGGCACCTCCGCAAGATCCTGGTGTGAGAGACTGTGACGAGCCCCGAGCC	1459
Qy	1441	ATCGCACTGACCCCGGTGATGGGCCTCATGTACTGGA CAGACTGGGAGAGAACCTTAAA	1500

Db 2540 GACACCAACATGATCGAGTCTGTCACCATGCTGGGTGAGGAGCGGGTCTGATTGCCGAC 2599
Qy 2581 GATCTCCCGCACCCGTTTCGGTCTGACGACGACGATACAGCGATTATATCTACTGACAGACTGG 2640
Db 2600 GATCTCCCGCACCCGTTTCGGTCTGACGACGATACAGCGATTATATCTACTGACAGACTGG 2659
Qy 2641 AATCTGCACAGCATTTGAGCGGCGCGACAAGACTAGCGCGCGAAACCGCACCCCTCATCCAG 2700
Db 2660 AATCTGCACAGCATTTGAGCGGCGCGACAAGACTAGCGGCGCGAAACCGCACCCCTCATCCAG 2719
Qy 2701 GGCACACTGGATTCTGATGAGACATCTGTGTTTCCACTCTCTCCGCGCAGATGGCCCTC 2760
Db 2720 GGCACACTGGACTCTGATGAGACATCTGTGTTTCCACTCTCTCCGCGCAGATGGCCCTC 2779
Qy 2761 AATGACTGTATGCAACAACGCGGAGGTGGGACGCTGTGCTTGGCCATCCCGCGCGC 2820
Db 2780 AATGACTGTATGCAACAACGCGGAGGTGGGACGCTGTGCTTGGCCATCCCGCGCGC 2839
Qy 2821 CACCGCTGCGGCTGCGCTCACACTACACCCCTGGACCCGACGCGCGCAACTGACGCGCG 2880
Db 2840 CACCGCTGCGGCTGCGCTCACACTACACCCCTGGACCCGACGCGCGCAACTGACGCGCG 2899
Qy 2881 CCCACCACTTCTGCTTTCAGCCAGAAATCTGCGATCATGTCGATGATCCCGGACGAC 2940
Db 2900 CCCACCACTTCTGCTTTCAGCCAGAAATCTGCGATCATGTCGATGATCCCGGACGAC 2959
Qy 2941 CAGCACGCGCGGATCTCATCTGCGCCCTGCATGACCTGAGGACGTCGAAAGCCATCGAC 3000
Db 2960 CAGCACGCGCGGATCTCATCTGCGCCCTGCATGACCTGAGGACGTCGAAAGCCATCGAC 3019
Qy 3001 TATGACCACTGGACAAAGTTTCATCTACTGCGGTGATGGGCGCCAGAAATCAAGCGAGCC 3060
Db 3020 TATGACCACTGGACAAAGTTTCATCTACTGCGGTGATGGGCGCCAGAAATCAAGCGAGCC 3079
Qy 3061 AAGGACGAGCGGACCCAGCCCTTTGTTTGTGACCTCTTGAGCCAAAGCCAAACCCAGAC 3120
Db 3080 AAGGACGAGCGGACCCAGCCCTTTGTTTGTGACCTCTTGAGCCAAAGCCAAACCCAGAC 3139
Qy 3121 AGGACGCGCCACGACCTCAGCATCGACATCTACAGCGGACACTGTTCTGACGTCGGAG 3180
Db 3140 AGGACGCGCCACGACCTCAGCATCGACATCTACAGCGGACACTGTTCTGACGTCGGAG 3199
Qy 3181 GCCACCAATACATCAACGCTCCACAGGCTGAGCGGGGAGCCATGGGGGTTGTTGCGT 3240
Db 3200 GCCACCAATACATCAACGCTCCACAGGCTGAGCGGGGAGCCATGGGGGTTGTTGCGT 3259
Qy 3241 GGGGACCGCGACAAGCCAGGGCCATCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 3300
Db 3260 GGGGACCGCGACAAGCCAGGGCCATCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 3319
Qy 3301 ACCAATGACGAGACCGGGACGCGAAGATCGAAGCGCGGACCTGGAGCGGACCGAGCGC 3360
Db 3320 ACCAATGACGAGACCGGGACGCGAAGATCGAAGCGCGGACCTGGAGCGGACCGAGCGC 3379
Qy 3361 GAGGTCTCTTTCACACCGGCTCATCCGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 3420
Db 3380 GAGGTCTCTTTCACACCGGCTCATCCGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 3439
Qy 3421 GGCAGGCTGTTCTGGGTGAGCGCGACCTGAGCGCATTTGAGAGCTGTGACCTGTGAGG 3480
Db 3440 GGCAGGCTGTTCTGGGTGAGCGCGACCTGAGCGCATTTGAGAGCTGTGACCTGTGAGG 3499
Qy 3481 GCCAACCGCTGACCTGGAGGAGCGCAACATCTGTGAGCCTCTGGGCTGAGCAATCTT 3540
Db 3500 GCCAACCGCTGACCTGGAGGAGCGCAACATCTGTGAGCCTCTGGGCTGAGCAATCTT 3559
Qy 3541 GGCAGGATCTCTACTGATGAGCGCGGACGAGGATGATCGAGCGGTGTGAGAGACG 3600
Db 3560 GGCAGGATCTCTACTGATGAGCGCGGACGAGGATGATCGAGCGGTGTGAGAGACG 3619
Qy 3601 ACCGGGACAAAGCGGACTTCGATCCAGGGCGGTGTGCGCCACCTCACTGGCATCCATGCA 3660
Db 3620 ACCGGGACAAAGCGGACTTCGATCCAGGGCGGTGTGCGCCACCTCACTGGCATCCATGCA 3679

Qy 3661 GTGAGGAAGTCAGCTGGAGAGTTTCTACGCCACCCATCTGTGCCGTGACAAATGCTGCGC 3720
Db 3680 GTGAGGAAGTCAGCTGGAGAGTTTCTACGCCACCCATCTGTGCCGTGACAAATGCTGCGC 3739
Qy 3721 TGCTCCCAATCTGTATTTGCCAAGGGTGTATGGGACACACAGGTGTCTATGCCCCAGTCCAC 3780
Db 3740 TGCTCCCAATCTGTATTTGCCAAGGGTGTATGGGACACACAGGTGTCTATGCCCCAGTCCAC 3799
Qy 3781 CTGCTGCTCTGACAACTGCTGACCTGTGAGAGCGCGCCACCTGTCTCCCGGACGACG 3840
Db 3800 CTGCTGCTCTGACAACTGCTGACCTGTGAGAGCGCGCCACCTGTCTCCCGGACGACG 3859
Qy 3841 TTTGCACTGTGACACAGGGGAGATCGACTGTATCCCGGGGCTGCGCTGTGTGACGCGCTTT 3900
Db 3860 TTTGCACTGTGACACAGGGGAGATCGACTGTATCCCGGGGCTGCGCTGTGTGACGCGCTTT 3919
Qy 3901 CCCAGTGCATGATGACGAGCGAGAGGGCTGCCCCGTGTGTCTCCCGCGCCAGTTTC 3960
Db 3920 CCCAGTGCATGATGACGAGCGAGAGGGCTGCCCCGTGTGTCTCCCGCGCCAGTTTC 3979
Qy 3961 CCTTGGCGCGGGTCAAGTGTGTGACCTGTGCGCTGTGCTGCGAGCGGAGCGAGACTGT 4020
Db 3980 CCTTGGCGCGGGTCAAGTGTGTGACCTGTGCGCTGTGCTGCGAGCGGAGCGAGACTGT 4039
Qy 4021 CAGGACCGCTCAGACGAGGGGACTGTGACGCGCATCTGCGTCCCAACAGTTCGCGTGT 4080
Db 4040 CAGGACCGCTCAGACGAGGGGACTGTGACGCGCATCTGCGTCCCAACAGTTCGCGTGT 4099
Qy 4081 GCGAGCGCCAGTGTCTCTCATCAAAACAGCAGTGCAGCTCTTCTCCCGACTGTATCGAC 4140
Db 4100 GCGAGCGCCAGTGTCTCTCATCAAAACAGCAGTGCAGCTCTTCTCCCGACTGTATCGAC 4159
Qy 4141 GGCTCCGACGAGCTCATGTGTGAAATCAACGAGCGGCTCTCAGACGACGCGCGCCAC 4200
Db 4160 GGCTCCGACGAGCTCATGTGTGAAATCAACGAGCGGCTCTCAGACGACGCGCGCCAC 4219
Qy 4201 AGCAGTGCATCGGCGCGCTCATTTGGGACTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4260
Db 4220 AGCAGTGCATCGGCGCGCTCATTTGGGACTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4279
Qy 4261 TATTTTGTGTCACGCGCTGTGCGCGCTATGCGGGGCGCAACGCGGCGCTTCCG 4320
Db 4280 TATTTTGTGTCACGCGCTGTGCGCGCTATGCGGGGCGCAACGCGGCGCTTCCG 4339
Qy 4321 CACGAGTATGTACGCGGACCCCCGACGCTGCGCTCTCAATTTTCATAGCCCCGCGGCTTCC 4380
Db 4340 CACGAGTATGTACGCGGACCCCCGACGCTGCGCTCTCAATTTTCATAGCCCCGCGGCTTCC 4399
Qy 4381 CAGCATGCGCCCTTTCACAGGATCGCATGCGGAAAGTCCATGATGAGCTCTCGTGAAGCTG 4440
Db 4400 CAGCATGCGCCCTTTCACAGGATCGCATGCGGAAAGTCCATGATGAGCTCTCGTGAAGCTG 4459
Qy 4441 ATGGGGGCGCGGGGCTGCCCTGTACGACCGGAAACACGTCACAGGGGCTCTCGTCC 4500
Db 4460 ATGGGGGCGCGGGGCTGCCCTGTACGACCGGAAACACGTCACAGGGGCTCTCGTCC 4519
Qy 4501 AGCAGTCTGTCCAGCAAGAGGCGACGCTGTATCCCGCGGATCTCTGAACCCCGCGCCCTCC 4560
Db 4520 AGCAGTCTGTCCAGCAAGAGGCGACGCTGTATCCCGCGGATCTCTGAACCCCGCGCCCTCC 4579
Qy 4561 CCGGCGACGAGCCCTCTCTGTACAAATGAGACTGTCTACTCTTCTTCTTCTTCTTCTTCTTCTTCT 4620
Db 4580 CCGGCGACGAGCCCTCTCTGTACAAATGAGACTGTCTACTCTTCTTCTTCTTCTTCTTCTTCTTCT 4639
Qy 4621 ACTCGGAGACGTCAGCGCCCTCATCATTTGAGAAATGGCGCGCCCGACGACGCGCTTCC 4680
Db 4640 ACTCGGAGACGTCAGCGCCCTCATCATTTGAGAAATGGCGCGCCCGACGACGCGCTTCC 4699
Qy 4681 AGCACCGAGTGTGTGACGAGACTACAGCGCCAGCGCTGTGAAGGCGCAGCAAGTACTTAC 4740
Db 4700 AGCACCGAGTGTGTGACGAGACTACAGCGCCAGCGCTGTGAAGGCGCAGCAAGTACTTAC 4759

Db 1400 GCGCTCAACGGCACCTCCGCAAGATCTGGTGTGCGAGGACCTGGACGAGCCCGAGCC 1459
QY 1441 ATCGACATGCAACCCGCTGATGGCCCTCATGTACTGGACAGACTGGGGAGAGAACCTTAAA 1500
Db 1460 ATCGACATGCAACCCGCTGATGGCCCTCATGTACTGGACAGACTGGGGAGAGAACCTTAAA 1519
QY 1501 ATCGAGTGTGCAACTTGGATGGGGAGGAGCGGCTGTGCTGGTCAATGCCCTCCCTCGGG 1560
Db 1520 ATCGAGTGTGCAACTTGGATGGGGAGGAGCGGCTGTGCTGGTCAATGCCCTCCCTCGGG 1579
QY 1561 TGGCCCAACGGCTGGCCCTGGACCTGCAGAGGGGAAAGCTCTACTGGGGAGAGCCCAAG 1620
Db 1580 TGGCCCAACGGCTGGCCCTGGACCTGCAGAGGGGAAAGCTCTACTGGGGAGAGCCCAAG 1639
QY 1621 ACAGACAAGATCGAGGTGATCAATGTTGATGGGACGAAGAGCGGACCCCTCTGGAGGAC 1680
Db 1640 ACAGACAAGATCGAGGTGATCAATGTTGATGGGACGAAGAGCGGACCCCTCTGGAGGAC 1699
QY 1681 AAGCTCCCGCAATTTTCGGGTTTCACTGCTGGGGGACTTCACTACTGGAAGTACTGG 1740
Db 1700 AAGCTCCCGCAATTTTCGGGTTTCACTGCTGGGGGACTTCACTACTGGAAGTACTGG 1759
QY 1741 CAGCGCCGACGATCGAGCGGGTGCACAAAGGTCAAGGCCAGCGGGAAGTCACTATTGAC 1800
Db 1760 CAGCGCCGACGATCGAGCGGGTGCACAAAGGTCAAGGCCAGCGGGAAGTCACTATTGAC 1819
QY 1801 CAGCTGCCCGACCTGATGGGGTCAAGCTGTGAATGTGGCCAAAGGTGCGTGGAAACCAAC 1860
Db 1820 CAGCTGCCCGACCTGATGGGGTCAAGCTGTGAATGTGGCCAAAGGTGCGTGGAAACCAAC 1879
QY 1861 CCGTGTGGGACAGAAACGGGGGTGACAGCCACTGTGCTTTCACACCCCGACGCAACC 1920
Db 1880 CCGTGTGGGACAGAAACGGGGGTGACAGCCACTGTGCTTTCACACCCCGACGCAACC 1939
QY 1921 CCGTGTGGCTGCCCATCGGCTTGAGCTGTGAGTGCATGAAGACCTGCTGCTGCT 1980
Db 1940 CCGTGTGGCTGCCCATCGGCTTGAGCTGTGAGTGCATGAAGACCTGCTGCTGCT 1999
QY 1981 GAGGCTTTCTTGGTCTTCAAGAGAGCGCGCATTCACAGAGTCTCCCTCGAGACCAAT 2040
Db 2000 GAGGCTTTCTTGGTCTTCAAGAGAGCGCGCATTCACAGAGTCTCCCTCGAGACCAAT 2059
QY 2041 AACACGAGCTGGCCATCCGCTCAGCGGCTCAAGGAGGCTCAGCCCTGAGCTTTGAT 2100
Db 2060 AACACGAGCTGGCCATCCGCTCAGCGGCTCAAGGAGGCTCAGCCCTGAGCTTTGAT 2119
QY 2101 GTGTCCAAACACCATCTACTGGAAGAGCTGAGCTGAGAGCCATCAGCGCGGCTTC 2160
Db 2120 GTGTCCAAACACCATCTACTGGAAGAGCTGAGCTGAGAGCCATCAGCGCGGCTTC 2179
QY 2161 ATGAACGGGAGCTCGGTGGAGCACTGTGTGAGTTTGGCCCTTGATACCCGAGGGCATG 2220
Db 2180 ATGAACGGGAGCTCGGTGGAGCACTGTGTGAGTTTGGCCCTTGATACCCGAGGGCATG 2239
QY 2221 GCCGTGATCTGATGGGCAAGAACTTACTTGGGCGGACACTGGGACCAACAGATCGAA 2280
Db 2240 GCCGTGATCTGATGGGCAAGAACTTACTTGGGCGGACACTGGGACCAACAGATCGAA 2299
QY 2281 GTGGCGGCTGGAGCGGCGAGTTCCGGCAAGTCTCTGTGTGGAGGAGCTTGGACAAACCCG 2340
Db 2300 GTGGCGGCTGGAGCGGCGAGTTCCGGCAAGTCTCTGTGTGGAGGAGCTTGGACAAACCCG 2359
QY 2341 AGTGTGCTGGCCCTGGATCCCAACGAGGCTTACTTACTGGAACGAGTGGGGCGGCAAG 2400
Db 2360 AGTGTGCTGGCCCTGGATCCCAACGAGGCTTACTTACTGGAACGAGTGGGGCGGCAAG 2419
QY 2401 CCGAGGATCTGGCGGCTTTCATGACCGGACCAACTGATGACGCTGGTGGACAAAGTG 2460
Db 2420 CCGAGGATCTGGCGGCTTTCATGACCGGACCAACTGATGACGCTGGTGGACAAAGTG 2479
QY 2461 GGCGGGGCAACGACCTCACCATTGACTACGCTGACGAGCGCTTCTACTGGACCGACCTG 2520

Db 2480 GGCGGGCCACGACCTTCAACATTGACTACGCTGACAGCGCCTTACTTGGACCGACCTG 2539
QY 2521 GACACCAACATGATCGAGTCTGTCACACATGCTGGGTGAGGAGCGGCTGCTGATGTCGAC 2580
Db 2540 GACACCAACATGATCGAGTCTGTCACACATGCTGGGTGAGGAGCGGCTGCTGATGTCGAC 2599
QY 2581 GATCTCCGCGACCCGTTTGGTCTGACGCACTGACGCAATATATCTACTGACAGACTGG 2640
Db 2600 GATCTCCGCGACCCGTTTGGTCTGACGCACTGACGCAATATATCTACTGACAGACTGG 2659
QY 2641 AATCTGCAACAGCATGTAGCGGCGGCGCAAGACTAGCGGCGGAAACCGACCTCATCCAG 2700
Db 2660 AATCTGCAACAGCATGTAGCGGCGGCGCAAGACTAGCGGCGGAAACCGACCTCATCCAG 2719
QY 2701 GGCCACCTGGAGTTTGGTATGAGCATCTCTGGTGTTCACCTCTCCCGCAGAGTGGCTC 2760
Db 2720 GGCCACCTGGAGTTTGGTATGAGCATCTCTGGTGTTCACCTCTCCCGCAGAGTGGCTC 2779
QY 2761 AATGACTGTATGCAACAAACGGGCGAGTGTGGGCACTGTGCTTGCCTATCCCGCGGCGC 2820
Db 2780 AATGACTGTATGCAACAAACGGGCGAGTGTGGGCACTGTGCTTGCCTATCCCGCGGCGC 2839
QY 2821 CACCGCTGCGGCTCGGCTTCACTACACCTTGACCCCGACGAGCCGCAACTGACGCGG 2880
Db 2840 CACCGCTGCGGCTCGGCTTCACTACACCTTGACCCCGACGAGCGCAACTGACGCGG 2899
QY 2881 CCAACGACCTTCTTGTGTTTTCAGCGCAAGAAATCTGCGCATCAGTCGAGTATCCGAGCAG 2940
Db 2900 CCAACGACCTTCTTGTGTTTTCAGCGCAAGAAATCTGCGCATCAGTCGAGTATCCGAGCAG 2959
QY 2941 CAGCAGCGCGGATCTCATCTCCCTGATGAGTGTGAGGAAAGTCAAGACCATCGAC 3000
Db 2960 CAGCAGCGCGGATCTCATCTCCCTGATGAGTGTGAGGAAAGTCAAGACCATCGAC 3019
QY 3001 TATGACCCACTGGAACAAGTTTCATCTACTGGTGGATGGCGGCCAGAAATCAAGCGAGCC 3060
Db 3020 TATGACCCACTGGAACAAGTTTCATCTACTGGTGGATGGCGGCCAGAAATCAAGCGAGCC 3079
QY 3061 AAGGACGAGCGGACCCAGCCCTTGTGTTTGTGACTCTCTGAGCCAAAGGCAAAACCCAGAC 3120
Db 3080 AAGGACGAGCGGACCCAGCCCTTGTGTTTGTGACTCTCTGAGCCAAAGGCAAAACCCAGAC 3139
QY 3121 AGGCGAGCCCAACGACCTCAGCATCGACATCTACAGCGGACACTGTTCTTGACGCTGCGAG 3180
Db 3140 AGGCGAGCCCAACGACCTCAGCATCGACATCTACAGCGGACACTGTTCTTGACGCTGCGAG 3199
QY 3181 GCCACCAATACCATCAACGCTCCAGAGCTGAGCGGGGAAGCCATGGGGGTGTTGCTGCT 3240
Db 3200 GCCACCAATACCATCAACGCTCCAGAGCTGAGCGGGGAAGCCATGGGGGTGTTGCTGCT 3259
QY 3241 GGGGACCGGACCAAGCCAGGGCCATCGTGTGTAACGGGAGCGAGGGTACTGTACTTTC 3300
Db 3260 GGGGACCGGACCAAGCCAGGGCCATCGTGTGTAACGGGAGCGAGGGTACTGTACTTTC 3319
QY 3301 ACCAAATGACGAGCCGGGAGGCAAGATCGAACGCGAGCCCTGGAGCGGACCCAGCGC 3360
Db 3320 ACCAAATGACGAGCCGGGAGGCAAGATCGAACGCGAGCCCTGGAGCGGACCCAGCGC 3379
QY 3361 GAGTCTCTTTCACACCGGCTCATCCGCCCTGTGCGCTTGGTGGTGGACCAACACTG 3420
Db 3380 GAGTCTCTTTCACACCGGCTCATCCGCCCTGTGCGCTTGGTGGTGGACCAACACTG 3439
QY 3421 GGAAGCTGTTCTGGGTGGAGCGGACCTGAGGCGCATTTGAGAGCTGTGACTGTACAGG 3480
Db 3440 GGAAGCTGTTCTGGGTGGAGCGGACCTGAGGCGCATTTGAGAGCTGTGACTGTACAGG 3499
QY 3481 GCCAACCGCTGACCTTGGAGGAGCGCAACATCTGTGAGCTCTTGGGCTTGCACCATCTT 3540
Db 3500 GCCAACCGCTGACCTTGGAGGAGCGCAACATCTGTGAGCTCTTGGGCTTGCACCATCTT 3559
QY 3541 GGCAGGATCTTCTACTGATCGACCGGCGAGCAGAGATGATCGAGCGCTGTGAGAGAGACC 3600
Db 3560 GGCAGGATCTTCTACTGATCGACCGGCGAGCAGAGATGATCGAGCGCTGTGAGAGAGACC 3619

Db 241 GCGCGCAGTGGACTTCCAGTGTTCACAGGAGCCGTGTACTGGACAGAGCTGAGCGAG 300
QY 304 GAGGCCATCAGACAGACTCTGNAACAGAGCGGGGCGCGGTGCAGAACTGGTTCATC 363
Db 301 GAGGCCATCAGACAGACTCTGNAACAGAGCGGGGCGCGGTGCAGAACTGGTTCATC 360
QY 364 TCCGGCCTGGTCTCTCCGACGGCTCGCCTCGACTGGGTGGGCAAGACTGTACTGG 423
Db 361 TCCGGCCTGGTCTCTCCGACGGCTCGCCTCGACTGGGTGGGCAAGACTGTACTGG 420
QY 424 ACAGACTCAGAGACCAACCCGATCGAGGTGSCAACCTCAATGACACATCCCGAAGGTG 483
Db 421 ACAGACTCAGAGACCAACCCGATCGAGGTGSCAACCTCAATGACACATCCCGAAGGTG 480
QY 484 CTCTTCTGGCAGGACCTTGACAGCCTAGGCGCATCGCCTTGGACCCCGCTCAGCGGTAC 543
Db 481 CTCTTCTGGCAGGACCTTGAACAGCCGAGGCGCATCGCCTTGGACCCCGCTCAGCGGTAC 540
QY 544 ATGTACTGGAAGACTGGGGTGAGACGCCCGGATTTGAGCGGCGAGGATGGATGGCAGC 603
Db 541 ATGTACTGGAAGACTGGGGTGAGACGCCCGGATTTGAGCGGCGAGGATGGATGGCAGC 600
QY 604 ACCCGGAAGATCAATGTGGAATCGGACATTTACTGGGCCAATGGAATGACCATCGACCTG 663
Db 601 ACCCGGAAGATCAATGTGGAATCGGACATTTACTGGGCCAATGGAATGACCATCGACCTG 660
QY 664 GAGGACGAGAGCTTACTGGGCTGAGCGCAAGCTCAGCTTCAATCCACCGTGCACACTG 723
Db 661 GAGGACGAGAGCTTACTGGGCTGAGCGCAAGCTCAGCTTCAATCCACCGTGCACACTG 720
QY 724 GACGGCTGTTCCGCGAAGGTGTGAGGCGAGCTGAGCGACCCCTTCGCGCTGACG 783
Db 721 GACGGCTGTTCCGCGAAGGTGTGAGGCGAGCTGAGCGACCCCTTCGCGCTGACG 780
QY 784 CTCTCCGGGGAACATCTGTACTGGAAGACTGGCAGACCCGCTCCATCCATGCTGCAAC 843
Db 781 CTCTCCGGGGAACATCTGTACTGGAAGACTGGCAGACCCGCTCCATCCATGCTGCAAC 840
QY 844 AAGCGCACTGGGGGGAAGAGAGATCTGTAGTGCCCTTACTCAATCCATGAGACATC 903
Db 841 AAGCGCACTGGGGGGAAGAGAGATCTGTAGTGCCCTTACTCAATCCATGAGACATC 900
QY 904 CAGGTGCTGAGCAGGAGCGCAGCTTCTTCCACACTCGCTGTGAGGAGGACAAATGGC 963
Db 901 CAGGTGCTGAGCAGGAGCGCAGCTTCTTCCACACTCGCTGTGAGGAGGACAAATGGC 960
QY 964 GGCTGCTCCACCTGTGCTGCTGTCCCAAGCGAGCCTTTCTACATGCGCCTGCCCC 1023
Db 961 GGCTGCTCCACCTGTGCTGCTGTCCCAAGCGAGCCTTTCTACACATGCGCCTGCCCC 1020
QY 1024 ACGGGTGTGAGCTGCAGGACAAACCGGAGGAGCTGTAGGCGAGGAGCGGAGGCTGCTG 1083
Db 1021 ACGGGTGTGAGCTGCAGGACAAACCGGAGGAGCTGTAGGCGAGGAGCGGAGGCTGCTG 1080
QY 1084 CTGCTGGCCCGCGGACGGAAGCTACGAGGAGTCTCGCTGGACACGCGGAGCTTCAACGAC 1143
Db 1081 CTGCTGGCCCGCGGACGGAAGCTACGAGGAGTCTCGCTGGACACGCGGAGCTTCAACGAC 1140
QY 1144 ATGCTGTGAGGTGAGGACATCCGGCAACCGGCAATGCGCATTCGACTACGACCCGCTAGAG 1203
Db 1141 ATGCTGTGAGGTGAGGACATCCGGCAACCGGCAATGCGCATTCGACTACGACCCGCTAGAG 1200
QY 1204 GGCTATGTCTACTGGAAGATGACGAGGTGGGGCCATCCGAGGCGGTACTCTGACGGG 1263
Db 1201 GGCTATGTCTACTGGAAGATGACGAGGTGGGGCCATCCGAGGCGGTACTCTGACGGG 1260
QY 1264 TCTGGGGGCGAGCGCTGCTCAAACCGGAGATCAACCGGAGTGGATGCGCTGCGCTGAC 1323
Db 1261 TCTGGGGGCGAGCGCTGCTCAAACCGGAGATCAACCGGAGTGGATGCGCTGCGCTGAC 1320
QY 1324 TGGGTGGCCGGAACCTCTACTTGGACCGACACGCGGACCGGATCAGGCTGAGCGCGC 1383

Db 1321 TGGGTGGCCGGAACCTCTACTGGACCGACACGCGGACCGGATCAGAGGTGACGCGC 1380
QY 1384 CTCAACGGCACCTCCCGAAGATCTGTGTGCGAGGAGCTGGACGAGCCCGAGCCATC 1443
Db 1381 CTCAACGGCACCTCCCGAAGATCTGTGTGCGAGGAGCTGGACGAGCCCGAGCCATC 1440
QY 1444 GCATCTGACACCCCGTGTATGGGCTCATGTACTGGACAGACTGGGGAGAGAACTTAAATC 1503
Db 1441 GCATCTGACACCCCGTGTATGGGCTCATGTACTGGACAGACTGGGGAGAGAACTTAAATC 1500
QY 1504 GAGTGTGCAAACTTGGATGGGAGGAGCGGCTGTGTGTGCTCAATGCTCTCCTCGGGTGG 1563
Db 1501 GAGTGTGCAAACTTGGATGGGAGGAGCGGCTGTGTGTGCTCAATGCTCTCCTCGGGTGG 1560
QY 1564 CCCAACGGCCTGGCCCTGGACCTCGACGAGGGGAGCTCTACTGGGGAGAGCCCAAGACA 1623
Db 1561 CCCAACGGCCTGGCCCTGGACCTCGACGAGGGGAGCTCTACTGGGGAGAGCCCAAGACA 1620
QY 1624 GACAAGATCGAGGTGATCAATGTTGATGGGACGAGAGCGGACCCCTCTCGAGGACAAAG 1683
Db 1621 GACAAGATCGAGGTGATCAATGTTGATGGGACGAGAGCGGACCCCTCTCGAGGACAAAG 1680
QY 1684 CTCCTGCAATTTTCGGGTTTCACTGCTGGGGGACTTTCATCTACTGGAATGACTGGCAG 1743
Db 1681 CTCCTGCAATTTTCGGGTTTCACTGCTGGGGGACTTTCATCTACTGGAATGACTGGCAG 1740
QY 1744 CGCGCAGCATCGAGCGGTCGACAGGTGCAAGGTCGAGCGGAGCTCATCATTCGACCAAG 1803
Db 1741 CGCGCAGCATCGAGCGGTCGACAGGTGCAAGGTCGAGCGGAGCTCATCATTCGACCAAG 1800
QY 1804 CTGCGCCGACCTGTGAGTGGGCTCAAAAGCTGTGATGTGGCCAAAGGTCGTCGGAACCAACCCG 1863
Db 1801 CTGCGCCGACCTGTGAGTGGGCTCAAAAGCTGTGATGTGGCCAAAGTCTGTCGGAACCAACCCG 1860
QY 1864 TGTGCGGACAGGAAACGGGGGTCGAGCCACTGTGCTTTCACACCCCAACGAAACCCG 1923
Db 1861 TGTGCGGACAGGAAACGGGGGTCGAGCCACTGTGCTTTCACACCCCAACGAAACCCG 1920
QY 1924 TGTGCTGCCCCATCGGCTCGAGCTGCTGAGTGCATGAGAGCTGCTGCTGCTGAG 1983
Db 1921 TGTGCTGCCCCATCGGCTCGAGCTGCTGAGTGCATGAGAGCTGCTGCTGCTGAG 1980
QY 1984 GCCTTCTTGTGCTTTCACGACGAGCGGCCATCCACAGGATCTCCCTCGAGACCAATTAAC 2043
Db 1981 GCCTTCTTGTGCTTTCACGACGAGCGGCCATCCACAGGATCTCCCTCGAGACCAATTAAC 2040
QY 2044 AACGAGTGGCCATCCGCTCAACGAGGCTCAAGGAGGCTCAGGCTTGGATGTG 2103
Db 2041 AACGAGTGGCCATCCGCTCAACGAGGCTCAAGGAGGCTCAGGCTTGGATGTG 2100
QY 2104 TCCAAACACCATCTACTGGACAGAGCTGAGCTGAGACCATCAGCGCGGCTTCATG 2163
Db 2101 TCCAAACACCATCTACTGGACAGAGCTGAGCTGAGACCATCAGCGCGGCTTCATG 2160
QY 2164 AACGGGAGCTCGGTGGAGCAGCTGTGGAGTTTGGCTTGTACTACCCCGAGGCAATGCC 2223
Db 2161 AACGGGAGCTCGGTGGAGCAGCTGTGGAGTTTGGCTTGTACTACCCCGAGGCAATGCC 2220
QY 2224 GTTGAATGGGCAAGAACTCTACTGGGCGGACACTGGGACCAACAGAAATCGAAGTG 2283
Db 2221 GTTGAATGGGCAAGAACTCTACTGGGCGGACACTGGGACCAACAGAAATCGAAGTG 2280
QY 2284 GCGGGCTTGGACGGGAGTTCCGCAAGTCTCGTGTGGAGGAGCTTGGACAAACCCGAGG 2343
Db 2281 GCGGGCTTGGACGGGAGTTCCGCAAGTCTCGTGTGGAGGAGCTTGGACAAACCCGAGG 2340
QY 2344 TCGCTGGCCCTTGGATCCCAACAGGGCTTACTACTTGGACCGGAGTGGGGCGGCAAGCG 2403
Db 2341 TCGCTGGCCCTTGGATCCCAACAGGGCTTACTACTTGGACCGGAGTGGGGCGGCAAGCG 2400
QY 2404 AGGATCGTGGGGCTTCAATGGAAGGAGCAAACTGTGCTGAGCTGCTGCTGGAAGGTGGGC 2463
Db 2401 AGGATCGTGGGGCTTCAATGGAAGGAGCAAACTGTGCTGAGCTGCTGCTGGAAGGTGGGC 2460

QY 2464 CGGGCAACGACCTCAACGATTGACATGACGCTGACGAGCGCTCTACTGAGCCGACCTGGAC 2523
DB |||||
2461 CGGGCAACGACCTCAACGATTGACATGACGCTGACGAGCGCTCTACTGAGCCGACCTGGAC 2520
QY |||||
DB |||||
2524 ACCAATCATGATCGAGTCGTCGACCAACATGCTGGGTGAGGAGCGGTGATGATGCGACGAT 2583
DB |||||
2521 ACCAATCATGATCGAGTCGTCGACCAACATGCTGGGTGAGGAGCGGTGATGATGCGACGAT 2580
QY |||||
DB |||||
2584 CTCGCGCACCCGTTTCGGTCTGACGAGTACAGCGATTATATCTACTGAGACACTGGAA 2643
DB |||||
2581 CTCGCGCACCCGTTTCGGTCTGACGAGTACAGCGATTATATCTACTGAGACACTGGAA 2640
QY |||||
DB |||||
2644 CTGCAAGATTGAGGGGCGACAGACTAGCGCGGAAACCGACCTCTATCAGAGGC 2703
DB |||||
2641 CTGCAAGATTGAGGGGCGACAGACTAGCGCGGAAACCGACCTCTATCAGAGGC 2700
QY |||||
DB |||||
2704 CACCTGGACTCTGATGAGACATCTCGGTGTTCCACTCTCTCCGCGAGGATGGCTCAAT 2763
DB |||||
2701 CACCTGGACTCTGATGAGACATCTCGGTGTTCCACTCTCTCCGCGAGGATGGCTCAAT 2760
QY |||||
DB |||||
2764 GACTGATGCAACAAACGCGGAGTGTGGCGAGCTGTGCTTGGCCATCCCGCGGCGCAC 2823
DB |||||
2761 GACTGATGCAACAAACGCGGAGTGTGGCGAGCTGTGCTTGGCCATCCCGCGGCGCAC 2820
QY |||||
DB |||||
2824 CGCTGGGCTGGCTCTCACTACACCTTGAGCCCGGAGCGGCAACTGAGCGCGGCC 2883
DB |||||
2821 CGCTGGGCTGGCTCTCACTACACCTTGAGCCCGGAGCGGCAACTGAGCGCGGCC 2880
QY |||||
DB |||||
2884 ACCACTCTCTGCTGTTCAGCCAGAAATCTGCCATCAGTCGATGATCCCGAGACACAG 2943
DB |||||
2881 ACCACTCTCTGCTGTTCAGCCAGAAATCTGCCATCAGTCGATGATCCCGAGACACAG 2940
QY |||||
DB |||||
2944 CACAGCCGGAATCTCATCTCTGCTGATGATGAGTGTGAGGAGCTCAAGCCATCGACTAT 3003
DB |||||
2941 CACAGCCGGAATCTCATCTCTGCTGATGATGAGTGTGAGGAGCTCAAGCCATCGACTAT 3000
QY |||||
DB |||||
3004 GACCACTGGACAGTTCATCTACTCTGGGTGATGGCGGCCAGAAATCAAGCGAGCCAG 3063
DB |||||
3001 GACCACTGGACAGTTCATCTACTCTGGGTGATGGCGGCCAGAAATCAAGCGAGCCAG 3060
QY |||||
DB |||||
3064 GACGAGGAGCCAGCCCTTTGTTTGTGACCTCTCTGAGCGAAGGCCAAACCCAGACAGG 3123
DB |||||
3061 GACGAGGAGCCAGCCCTTTGTTTGTGACCTCTCTGAGCGAAGGCCAAACCCAGACAGG 3120
QY |||||
DB |||||
3124 CAGCCCGAGGACTCAGCATCGACATCTACAGCGGACACTGTTCTGAGAGTGCAGGCC 3183
DB |||||
3121 CAGCCCGAGGACTCAGCATCGACATCTACAGCGGACACTGTTCTGAGAGTGCAGGCC 3180
QY |||||
DB |||||
3184 ACCAATACCATCAACGTCACAGGCTGAGCGGGGAAGCCATGGGGTGTGCTGCTGGG 3243
DB |||||
3181 ACCAATACCATCAACGTCACAGGCTGAGCGGGGAAGCCATGGGGTGTGCTGCTGGG 3240
QY |||||
DB |||||
3244 GACCGGCAAGCCAGGGCCATCTGCTGCAACGCGGAGGAGGTACTGTTACTTACC 3303
DB |||||
3241 GACCGGCAAGCCAGGGCCATCTGCTGCAACGCGGAGGAGGTACTGTTACTTACC 3300
QY |||||
DB |||||
3304 AACATGAGGACCGGGAGCCAGATCGAAACGCGAGCCCTGAGCGGACCGAGCGCCAG 3363
DB |||||
3301 AACATGAGGACCGGGAGCCAGATCGAAACGCGAGCCCTGAGCGGACCGAGCGCCAG 3360
QY |||||
DB |||||
3364 GTCTCTTTTACCACCGGCTCATCTCCGCTCTGCGCTGTGGTGGTGGACAAACACACTGGGC 3423
DB |||||
3361 GTCTCTTTTACCACCGGCTCATCTCCGCTGTGGCTGTGGTGGTGGTGGTGGTGGG 3420
QY |||||
DB |||||
3424 AAGCTGTTCTGGTGGAGCGGGAACCTGAAGCGCATTTGAGAGCTGTGAACCTGTGAGGGCC 3483
DB |||||
3421 AAGCTGTTCTGGTGGAGCGGGAACCTGAAGCGCATTTGAGAGCTGTGAACCTGTGAGGGCC 3480
QY |||||
DB |||||
3484 AACGSCCTGACCTGGAGAGCGCAACATCGTGGAGCTCTGGGCTGACCATCTCTGGC 3543
DB |||||
3481 AACGSCCTGACCTGGAGAGCGCAACATCGTGGAGCTCTGGGCTGACCATCTCTGGC 3540

QY 3544 AAGCATCTTACTGATCGAACCGCCAGCAGAGATGATCGAGCGTGTGGAGAACACACC 3603
DB |||||
3541 AAGCATCTTACTGATCGAACCGCCAGCAGAGATGATCGAGCGTGTGGAGAACACACC 3600
QY |||||
DB |||||
3604 GGGGAACAGCGGACTCGCATCCAGGGCGTGTGCGCCACCTCACTGCGCATCCATCGAGTG 3663
DB |||||
3601 GGGGAACAGCGGACTCGCATCCAGGGCGTGTGCGCCACCTCACTGCGCATCCATCGAGTG 3660
QY |||||
DB |||||
3664 GAGGAAGTCAGCCTGGAGGAGTTCCTCAGCCACCCATGTGCGCGTGACAAATGGTGGCTGC 3723
DB |||||
3661 GAGGAAGTCAGCCTGGAGGAGTTCCTCAGCCACCCATGTGCGCGTGACAAATGGTGGCTGC 3720
QY |||||
DB |||||
3724 TCCCAATCTGTATTGCCAAGGTGATGGGACACACCGGTGTATGCGCCAGTGCACCTC 3783
DB |||||
3721 TCCCAATCTGTATTGCCAAGGTGATGGGACACACCGGTGTATGCGCCAGTGCACCTC 3780
QY |||||
DB |||||
3784 GTGCTCTCGAGAACCTCTGACCTGTGAGAGCGCGCCACCTGCTCCCGGACAGTTC 3843
DB |||||
3781 GTGCTCTCGAGAACCTCTGACCTGTGAGAGCGCGCCACCTGCTCCCGGACAGTTC 3840
QY |||||
DB |||||
3844 GCATGTGCCACAGGGAGATCGACTGTATCCCGGGGCTGTGCGCTGTGAOGGCTTTCCC 3903
DB |||||
3841 GCATGTGCCACAGGGAGATCGACTGTATCCCGGGGCTGTGCGCTGTGAOGGCTTTCCC 3900
QY |||||
DB |||||
3904 GAGTCCGATGACAGAGCGAGGAGGCTGCCCGTGTGCTCCGCGCCAGTTCGCC 3963
DB |||||
3901 GAGTCCGATGACAGAGCGAGGAGGCTGCCCGTGTGCTCCGCGCCAGTTCGCC 3960
QY |||||
DB |||||
3964 TGGCGCGGGTCACTGTGTGGAACCTGCGCTGCGCTGCGAGCGGAGGAGTGTGAG 4023
DB |||||
3961 TGGCGCGGGTCACTGTGTGGAACCTGCGCTGCGCTGCGAGCGGAGGAGTGTGAG 4020
QY |||||
DB |||||
4024 GACGCTCAGACGAGGCGGACTGTGAACCCATCTGCTGCGCCAAACAGTTCGGTGTGG 4083
DB |||||
4021 GACGCTCAGACGAGGCGGACTGTGAACCCATCTGCTGCGCCAAACAGTTCGGTGTGG 4080
QY |||||
DB |||||
4084 AGCGCCAGTGTCTCATCAACAGCAGTGTGCTCTTCCCGGACTGTATCGACGCG 4143
DB |||||
4081 AGCGCCAGTGTCTCATCAACAGCAGTGTGCTCTTCCCGGACTGTATCGACGCG 4140
QY |||||
DB |||||
4144 TCCGACGAGCTCATGTGTGAATCAACAGCGCGCTCAGACGAGCCCGGCCACAGC 4203
DB |||||
4141 TCCGACGAGCTCATGTGTGAATCAACAGCGCGCTCAGACGAGCCCGGCCACAGC 4200
QY |||||
DB |||||
4204 AGTGCATCGGGCGGCTGAGTGTGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4263
DB |||||
4201 AGTGCATCGGGCGGCTGAGTGTGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4260
QY |||||
DB |||||
4264 TTTGTGTCAGCGCGTGTGTCAGCGCTATGCGGGGCAACGGGCGCTTCCCGCAC 4323
DB |||||
4261 TTTGTGTCAGCGCGTGTGTCAGCGCTATGCGGGGCAACGGGCGCTTCCCGCAC 4320
QY |||||
DB |||||
4324 GAGTATGTCAGCGGAGCCCGCACTGTCCTCTCAATTTCTATGACCCCGGGCGGTTCCAG 4383
DB |||||
4321 GAGTATGTCAGCGGAGCCCGCACTGTCCTCTCAATTTCTATGACCCCGGGCGGTTCCAG 4380
QY |||||
DB |||||
4384 CATGGCCCTTTCAGAGGATCGCATGGGAAAGTTCATGATGAGCTCCGTGAGCTGTGATG 4443
DB |||||
4381 CATGGCCCTTTCAGAGGATCGCATGGGAAAGTTCATGATGAGCTCCGTGAGCTGTGATG 4440
QY |||||
DB |||||
4444 GGGGGCGGGGCGGGGCTCCCTGTAGACCGGAAACCACTGTCACAGGGGCTCTGTCACAGC 4503
DB |||||
4441 GGGGGCGGGGCGGGGCTCCCTCTACGACCGGAAACCACTGTCACAGGGGCTCTGTCACAGC 4500
QY |||||
DB |||||
4504 AGTCTGTCAGACGAGAGGCGACGCTGTACCGCGGATTCCTGAAACCGCGCGCTTCCCG 4563
DB |||||
4501 AGTCTGTCAGACGAGAGGCGACGCTGTACCGCGGATTCCTGAAACCGCGCGCTTCCCG 4560
QY |||||
DB |||||
4564 GCCACGAGCCCTCTCTGTCAACATGGAGATGTTCTACTCTTCAAAACATTCGCGGCACT 4623
DB |||||
4561 GCCACGAGCCCTCTCTGTCAACATGGAGATGTTCTACTCTTCAAAATTCGCGGCACT 4620
QY |||||
DB |||||
4624 GCGAGACGCTACAGGCGCTCATCATTCAGAGAAATGGCGCGCCCGGACGACGACGCTTCGAGC 4683

```
Db 4621 GTGAGACCGGTGAGAGCGGCTTACATCTTGGAGAAATGGCGCCCGGAGCGGCTTGCAGC 4680
Qy 4684 ACCGACGCTGTGTGAGAGCGGCTTACAGCGCGCGCTGGAAGCGGAGGAGTACTACCTG 4743
Db 4681 ACCGACGCTGTGTGAGAGCGGCTTACAGCGCGCGCTGGAAGCGGAGGAGTACTACCTG 4740
Qy 4744 GATTTGAACTCGGAGTACAGACCCCTTATCCACCCCGGAGGAGGAGTACTACCTG 4803
Db 4741 GATTTGAACTCGGAGTACAGACCCCTTATCCACCCCGGAGGAGGAGTACTACCTG 4800
Qy 4804 TCGGCGGAGGAGGAGTACTACGCGCGCTTGCAGCGGAGGAGTACTACCTGCTTTC 4863
Db 4801 TCGGCGGAGGAGGAGTACTACGCGCGCTTGCAGCGGAGGAGTACTACCTGCTTTC 4860
Qy 4864 CGGCGCGCTTGCAGCGGAGTACTACGCGCGGAGGAGTACTACCTGCGGCGGAGTACT 4923
Db 4861 CGGCGCGCTTGCAGCGGAGTACTACGCGCGGAGGAGTACTACCTGCGGCGGAGTACT 4920
Qy 4924 TCTGTGCCCTGTAATAGTTTAAATATGAACAAAGAAAATAATATTTATGATTTA 4983
Db 4921 TCTGTGCCCTGTAATAGTTTAAATATGAACAAAGAAAATAATATTTATGATTTA 4980
Qy 4984 AAAATAAATAATAGTTTAAATATGAACAAAGAAAATAATATTTATGATTTA 5043
Db 4981 AAAATAAATAATAGTTTAAATATGAACAAAGAAAATAATATTTATGATTTA 5040
Qy 5044 AGGGCTGGGAGGAGTACTTGTGA 5063
Db 5041 AGGGCTGGGAGGAGTACTTGTGA 5060
```

```
RESULT 11
LOCUS AF064548 5100 bp mRNA linear PRI 22-SEP-1998
DEFINITION Homo sapiens low-density lipoprotein receptor-related protein 5
(LRP5) mRNA, complete cds.
ACCESSION AF064548
VERSION AF064548.1 GI:3641526
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 5100)
AUTHORS Hey, P.J., Twells, R.C., Phillips, M.S., Yusuke, Nakagawa, Brown, S.D.,
Kawaguchi, Y., Cox, R., Guochun, Xie., Dugan, V., Hammond, H.,
Metzker, M.L., Todd, J.A. and Hess, J.F.
Cloning of a novel member of the low-density lipoprotein receptor
family
JOURNAL Gene 216 (1), 103-111 (1998)
MEDLINE 98382578
PUBMED 9714764
REFERENCE 2 (bases 1 to 5100)
AUTHORS Hey, P.J., Twells, R.C., Phillips, M.S., Nakagawa, Y., Brown, S.D.,
Kawaguchi, Y., Cox, R., Xie, G., Dugan, V., Hammond, H., Metzker, M.L.,
Todd, J.A. and Hess, J.F.
Direct Submission
JOURNAL Submitted (11-MAY-1998) Human Genetics, Merck Research Lab,
Sunnytown Pike, West Point, PA 19486, USA
FEATURES
source
1..5100
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q13"
1..5100
/gene="LRP5"
49..4896
/gene="LRP5"
/codon_start=1
/product="low-density lipoprotein receptor-related protein
```

QY 544 ATGTACTGGACAGACTGGGGTGAAGCGCCCGGATTGAGCGGGCAGAGGATGGATGGGACG 603
DB 541 ATGTACTGGACAGACTGGGGTGAAGCGCCCGGATTGAGCGGGCAGAGGATGGATGGGACG 600
QY 604 ACCCGGAGATCATTTGTGGACTCGGACATTTACTTGGCCCAATGGACTGACCATGGACCTG 663
DB 601 ACCCGGAGATCATTTGTGGACTCGGACATTTACTTGGCCCAATGGACTGACCATGGACCTG 660
QY 664 GAGGAGCAGAAAGCTCTACTGGGCTGACGCCAAGCTCAGCTTCATCCACCGTGCCAAACCTG 723
DB 661 GAGGAGCAGAAAGCTCTACTGGGCTGACGCCAAGCTCAGCTTCATCCACCGTGCCAAACCTG 720
QY 724 GACGGCTCGTTCCGGCAGAAAGTGTGTGAGGAGAGCTTGACGACCCCTTCCGCTTGAAG 783
DB 721 GACGGCTCGTTCCGGCAGAAAGTGTGTGAGGAGAGCTTGACGACCCCTTCCGCTTGAAG 780
QY 784 CTCTCCGGGACACTCTGTACTGGACAGACTGGCAGACCCGCTCCATCCATGCGCTGCAAC 843
DB 781 CTCTCCGGGACACTCTGTACTGGACAGACTGGCAGACCCGCTCCATCCATGCGCTGCAAC 840
QY 844 AAGCCCACTGGGGGAGAGGAGAGATCTGTAGTGCCCTCTACTCAACCCATGGACATC 903
DB 841 AAGCCCACTGGGGGAGAGGAGAGATCTGTAGTGCCCTCTACTCAACCCATGGACATC 900
QY 904 CAGGTGTGAGCAGGAGCGGAGCGCTTTCTTCCACACTCGTGTGAGGAGACAAATGGC 963
DB 901 CAGGTGTGAGCAGGAGCGGAGCGCTTTCTTCCACACTCGTGTGAGGAGACAAATGGC 960
QY 964 GGTGCTCCCACTGTGCTGTGTCGCCAAGCGAGCGCTTTCTACACATGCGCTGCCCC 1023
DB 961 GGTGCTCCCACTGTGCTGTGTCGCCAAGCGAGCGCTTTCTACATGCGCTGCCCC 1020
QY 1024 ACGGTGTGCACTGCAAGCAAAACGCGAGGAGCTGTAAAGCAGGAGCGCGAGAGTGCTG 1083
DB 1021 ACGGTGTGCACTGCAAGCAAAACGCGAGGAGCTGTAAAGCAGGAGCGCGAGAGTGCTG 1080
QY 1084 CTGCTGGCCCGGCGACGGAACCTACGAGAGATCTCGCTGGACACCGCCGAGCTTACCGAC 1143
DB 1081 CTGCTGGCCCGGCGACGGAACCTACGAGAGATCTCGCTGGACACCGCCGAGCTTACCGAC 1140
QY 1144 ATCTGTGCTGAGGTGGAACGACATCCGCGACGCAATGGCCATCGACTACGACCCGCTAGAG 1203
DB 1141 ATCTGTGCTGAGGTGGAACGACATCCGCGACGCAATGGCCATCGACTACGACCCGCTAGAG 1200
QY 1204 GGTATGTCTACTGGACAGATGACGAGGTGGGGCCATCCGCGAGGGCGTACTGGAACGGG 1263
DB 1201 GGTATGTCTACTGGACAGATGACGAGGTGGGGCCATCCGCGAGGGCGTACTGGAACGGG 1260
QY 1264 TCTGGGGCGCAGACGCTGGTCAACCCGAGATCAACGACCCCGATGGCATCGCGTGCAC 1323
DB 1261 TCTGGGGCGCAGACGCTGGTCAACCCGAGATCAACGACCCCGATGGCATCGCGTGCAC 1320
QY 1324 TGGGTGGCCGCAAACTCTACTGGACCGACCGGCAACCGCATCGAGGTGACGGCG 1383
DB 1321 TGGGTGGCCGCAAACTCTACTGGACCGGCAACCGCATCGAGGTGACGGCG 1380
QY 1384 CTCAACGGCACTCCCGCAAGATCTGTGTGTCGAGGACCTGGAACGAGCCCGATC 1443
DB 1381 CTCAACGGCACTCCCGCAAGATCTGTGTGTCGAGGACCTGGAACGAGCCCGATC 1440
QY 1444 GCACTGCAACCCGCTGATGGGCTCATGTACTGGACAGACTGGGGAGAGAACCTTAAATC 1503
DB 1441 GCACTGCAACCCGCTGATGGGCTCATGTACTGGACAGACTGGGGAGAGAACCTTAAATC 1500
QY 1504 GAGTGTGCCAACTTGGATGGGAGGAGCGCGTGTGTGCTGCTCAATGCGCTCCCTCGGGTGG 1563
DB 1501 GAGTGTGCCAACTTGGATGGGAGGAGCGCGTGTGTGCTGCTCAATGCGCTCCCTCGGGTGG 1560
QY 1564 CCGAACGGCTCGGCTTGAACCTGACGAGGAGGAGAGCTTACTTGGGGAGACCGCAAGACA 1623
DB 1561 CCGAACGGCTCGGCTTGAACCTGACGAGGAGGAGAGCTTACTTGGGGAGACCGCAAGACA 1620

QY 1624 GACAAGATCGAGGTGATCAATGTTGATGGGACGAGGCGGACCTCTCTGGAGGACAAG 1683
DB 1621 GACAAGATCGAGGTGATCAATGTTGATGGGACGAGGCGGACCTCTCTGGAGGACAAG 1680
QY 1684 CTCCCGCACATTTTCGGGTTTCACTGCTGGGGACTTCTACTACTGGACTGACTGGCAG 1743
DB 1681 CTCCCGCACATTTTCGGGTTTCACTGCTGGGGACTTCTACTACTGGACTGACTGGCAG 1740
QY 1744 CGCCGACAGCATCGAGCGGGTGCAAAAGGTCAAGGCCACCGGGGAGCTCATCATTTGACAG 1803
DB 1741 CGCCGACAGCATCGAGCGGGTGCAAAAGGTCAAGGCCACCGGGGAGCTCATCATTTGACAG 1800
QY 1804 CTGCCCGACCTGATGGGGCTCAAAAGCTGTGAATGTGGCCAAAGGTCTGCGGAACCAACCG 1863
DB 1801 CTGCCCGACCTGATGGGGCTCAAAAGCTGTGAATGTGGCCAAAGGTCTGCGGAACCAACCG 1860
QY 1864 TGTGCGGACAGAAAGCGGGGTGAGCCACCTGTGCTTTTTCACACCCACGCAACCCGG 1923
DB 1861 TGTGCGGACAGAAAGCGGGGTGAGCCACCTGTGCTTTTTCACACCCACGCAACCCGG 1920
QY 1924 TGTGGCTGCCCATCGGCTGGAGCTGTGTGATGACATGAAGACCTGTGATCGTCTGAG 1983
DB 1921 TGTGGCTGCCCATCGGCTGGAGCTGTGTGATGACATGAAGACCTGTGATCGTCTGAG 1980
QY 1984 GCCTTCTGTGTTTTCACAGCAGAGCGCCATCCACAGGATCTCCCTCGAGACCAATAAC 2043
DB 1981 GCCTTCTGTGTTTTCACAGCAGAGCGCCATCCACAGGATCTCCCTCGAGACCAATAAC 2040
QY 2044 AACGACGTGGCCATCCGCTCACGGGCGTCAAGGAGGCTCAGCCCTGGACTTTGATGTG 2103
DB 2041 AACGACGTGGCCATCCGCTCACGGGCGTCAAGGAGGCTCAGCCCTGGACTTTGATGTG 2100
QY 2104 TCCAAACACACATCTACTGACAGACGTCAGCTCTGAAGACCATCAGCCGCGCTTCATG 2163
DB 2101 TCCAAACACACATCTACTGACAGACGTCAGCTCTGAAGACCATCAGCCGCGCTTCATG 2160
QY 2164 AACGGGAGCTCGGTGGAGCAGCTGTGGAGTTTGGCCCTTGACTACCCCGAGGGATGCCC 2223
DB 2161 AACGGGAGCTCGGTGGAGCAGCTGTGGAGTTTGGCCCTTGACTACCCCGAGGGATGCCC 2220
QY 2224 GTTGACTGGATGGGCAAGAACCTCTACTGGGCGGACACTGGGACCAACAGAAATCGAAGTG 2283
DB 2221 GTTGACTGGATGGGCAAGAACCTCTACTGGGCGGACACTGGGACCAACAGAAATCGAAGTG 2280
QY 2284 GCGCGCTGGAGCGGCGAGTTCCGCGCAAGTCTCGTGTGGAGGAGCTTGGACAAACCCGAGG 2343
DB 2281 GCGCGCTGGAGCGGCGAGTTCCGCGCAAGTCTCGTGTGGAGGAGCTTGGACAAACCCGAGG 2340
QY 2344 TCGCTGGCCCTGGATCCCAACAGGGCTACATCTACTGGACCGAGTGGGGCGGCAAGCCG 2403
DB 2341 TCGCTGGCCCTGGATCCCAACAGGGCTACATCTACTGGACCGAGTGGGGCGGCAAGCCG 2400
QY 2404 AGGATCGTGGCGGCTTTCATGGACGGGACCAACTGTCATGAGCTGTGGTGGACAGGTGGGC 2463
DB 2401 AGGATCGTGGCGGCTTTCATGGACGGGACCAACTGTCATGAGCTGTGGTGGACAGGTGGGC 2460
QY 2464 GGGCCCAACGACCTTCAACATTTGACTAGCTGACGAGCCCTCTACTGGACCGACCTGAC 2523
DB 2461 GGGCCCAACGACCTTCAACATTTGACTAGCTGACGAGCCCTCTACTGGACCGACCTGAC 2520
QY 2524 ACCAACATGATCGAGTCCGTCACCAATGCTGGGTGAGGAGCGGGTGTGATTTGCCGACAT 2583
DB 2521 ACCAACATGATCGAGTCCGTCACCAATGCTGGGTGAGGAGCGGGTGTGATTTGCCGACAT 2580
QY 2584 CTCCGCGACCCGTTGCTGACGAGGATGACGAGATTTATCTACTTGAACAGACTGGAT 2643
DB 2581 CTCCGCGACCCGTTGCTGACGAGGATGACGAGATTTATCTACTTGAACAGACTGGAT 2640
QY 2644 CTGCAACGATTTGAGCGGGCGGACAAAGACTAGCGGCCGGAACCGCACCTCATCCAGGGC 2703
DB 2641 CTGCAACGATTTGAGCGGGCGGACAAAGACTAGCGGCCGGAACCGCACCTCATCCAGGGC 2700
QY 2704 CACCTGGGACTTCTGTGATGGACATCTGTGTGTTCACCTCTCCCGGACGAGTGGCCTCAAT 2763

Db 1561 CCCAACGGCCTGGCCCTGGAACCTGACGAGGGGAGCTCTACTGGGGAGACGCCAAGACA 1620
Qy 1624 GACAAAGATCGAGGTGATCAATGTTGATCGGACGAGAGGCGGACCTCTCGAGGACAAAG 1683
Db 1621 GACAAAGATCGAGGTGATCAATGTTGATCGGACGAGAGGCGGACCTCTCGAGGACAAAG 1680
Qy 1684 CTCGCGCACATTTTCGGGTTTCACTGCTGGGGGACTTCTACTAGACTGACTGGCAG 1743
Db 1681 CTCGCGCACATTTTCGGGTTTCACTGCTGGGGGACTTCTACTAGACTGACTGGCAG 1740
Qy 1744 CGCGCAGCATCGAGCGGGTGCACAAAGTCAAGGCCAGCGGAGCTCATATTGACCGAG 1803
Db 1741 CGCGCAGCATCGAGCGGGTGCACAAAGTCAAGGCCAGCGGAGCTCATATTGACCGAG 1800
Qy 1804 CTGCCCGACCTGATGGGGCTCAAGCTGTGAATGTGGCCAAAGTCTCGGAAACCAACCCG 1863
Db 1801 CTGCCCGACCTGATGGGGCTCAAGCTGTGAATGTGGCCAAAGTCTCGGAAACCAACCCG 1860
Qy 1864 TGTGCGGACAGAAACGGGGGTGCAGCCACTGTGCTCTTTTCAACCCCAACCGAACCCGG 1923
Db 1861 TGTGCGGACAGAAACGGGGGTGCAGCCACTGTGCTCTTTTCAACCCCAACCGAACCCGG 1920
Qy 1924 TGTGGCTGCCCATCGGCTCGAGCTGCTGAGTGAATGAAGACCTGATCGTGGCTGAG 1983
Db 1921 TGTGGCTGCCCATCGGCTCGAGCTGCTGAGTGAATGAAGACCTGATCGTGGCTGAG 1980
Qy 1984 GCCTTTCTGTGCTTCAACAGCAGAGCGCGCCATCCACAGGATCTCCCTCGAGACCAATAAC 2043
Db 1981 GCCTTTCTGTGCTTCAACAGCAGAGCGCGCCATCCACAGGATCTCCCTCGAGACCAATAAC 2040
Qy 2044 AACGACGTGGGCATCCCGCTCAACGGGGTCAAGGAGGCTCAGCGCTGGACTTTGATGTG 2103
Db 2041 AACGACGT-GCCATCCCGCTCAACGGGGTCAAGGAGGCTCAGCGCTGGACTTTGATGTG 2099
Qy 2104 TCCAAACCAACATCTACTGACAGAGCTGAGCTGAAGACCAATCAGCGCGGCTTCATG 2163
Db 2100 TCCAAACCAACATCTACTGACAGAGCTGAGCTGAAGACCAATCAGCGCGGCTTCATG 2159
Qy 2164 AACGGAGCTCGGTGGAGCAGTGTGGAGTTGGCTTTGACTACCCGAGGGGATGGCC 2223
Db 2160 AACGGAGCTCGGTGGAGCAGTGTGGAGTTGGCTTTGACTACCCGAGGGGATGGCC 2219
Qy 2224 GTTGACTGGATGGGCAAGAACTTACTGGGCGGACACTGGGACCAACAGAAATCGAAGTG 2283
Db 2220 GTTGACTGGATGGGCAAGAACTTACTGGGCGGACACTGGGACCAACAGAAATCGAAGTG 2279
Qy 2284 GCGGGCTGGAACGGGAGTTCGGCAAGTCTCGTGTGGAGGACTTGGACAAACCCGAGG 2343
Db 2280 GCGGGCTGGAACGGGAGTTCGGCAAGTCTCGTGTGGAGGACTTGGACAAACCCGAGG 2339
Qy 2344 TCGCTGGCCCTGGATCCCAACAGGGCTACATCTACTGGACCGAGTGGGGCGGCAAGCCG 2403
Db 2340 TCGCTGGCCCTGGATCCCAACAGGGCTACATCTACTGGACCGAGTGGGGCGGCAAGCCG 2399
Qy 2404 AGGATCGTGGGGCTTTCATGGACGGGACCAACTGCATGACGCTGGTGGACAAAGTGGGC 2463
Db 2400 AGGATCGTGGGGCTTTCATGGACGGGACCAACTGCATGACGCTGGTGGACAAAGTGGGC 2459
Qy 2464 CGGGCAACAGACCTCAACATTTGACTACGCTGACCGAGTATATCTACTGGACAGACTGGAAT 2523
Db 2460 CGGGCAACAGACCTCAACATTTGACTACGCTGACCGAGTATATCTACTGGACAGACTGGAAT 2519
Qy 2524 ACCAATGATCGAGTCTGCAATGCTGGGTGAGGAGCGGCTGATGTCGCGACGAT 2583
Db 2520 ACCAATGATCGAGTCTGCAATGCTGGGTGAGGAGCGGCTGATGTCGCGACGAT 2579
Qy 2584 CTCGCGCACCCGTTCCGCTGACGAGTACAGCGATTTATCTACTGGACAGACTGGAAT 2643
Db 2580 CTCGCGCACCCGTTCCGCTGACGAGTACAGCGATTTATCTACTGGACAGACTGGAAT 2639
Qy 2644 CTGCAACAGATTTGAGCGGGCGGACAAAGACTAGCGGCGGAAACCGCACCTCATCCAGGGC 2703
Db 2640 CTGCAACAGATTTGAGCGGGCGGACAAAGACTAGCGGCGGAAACCGCACCTCATCCAGGGC 2699

Qy 2704 CACCTGACATTTCTGTGATGAGCATCTCTGTGTGTTCCATCTCTCCCGCCAGGATGGCTCAAT 2763
Db 2700 CACCTGACATTTCTGTGATGAGCATCTCTGTGTGTTCCATCTCTCCCGCCAGGATGGCTCAAT 2759
Qy 2764 GACTGTATGACAAACAAACGGGCACTGTGGGAGCTGTGCTTGGCATCTCCCGGGGGCCAC 2823
Db 2760 GACTGTATGACAAACAAACGGGCACTGTGGGAGCTGTGCTTGGCATCTCCCGGGGGCCAC 2819
Qy 2824 CGCTGCGGCTGCGCTCACACTACACCTTGAGACCCCGAGCGGCAACTGAGCGCCGCC 2883
Db 2820 CGCTGCGGCTGCGCTCACACTACACCTTGAGACCCCGAGCGGCAACTGAGCGCCGCC 2879
Qy 2884 ACCACCTTTCTGTGTTTTCAGCCAGAAATCTGCCATCAGTTCGATGATCTCCCGGAGACACAG 2943
Db 2880 ACCACCTTTCTGTGTTTTCAGCCAGAAATCTGCCATCAGTTCGATGATCTCCCGGAGACACAG 2939
Qy 2944 CACAGCCGGATCTCATCTCTGCTTGCATGGAATGAGGAAACGTTCAAGCCATCGACTAT 3003
Db 2940 CACAGCCGGATCTCATCTCTGCTTGCATGGAATGAGGAAACGTTCAAGCCATCGACTAT 2999
Qy 3004 GACCCACTGAGCAAGTTTCTACTGCTGATGGGCGGCAAGACATCAAGCGAGCCCAAG 3063
Db 3000 GACCCACTGAGCAAGTTTCTACTGCTGATGGGCGGCAAGACATCAAGCGAGCCCAAG 3059
Qy 3064 GACGACGGGACCCAGCCCTTTGTTTTCAGCTCTCTGAGCCAAAGGCCAAACCCAGACAGG 3123
Db 3060 GACGACGGGACCCAGCCCTTTGTTTTCAGCTCTCTGAGCCAAAGGCCAAACCCAGACAGG 3119
Qy 3124 CAGCCCAACGACTCAGCATCGACATCTACAGCGGACACTGTGTTTGGACGTTGGAGGCC 3183
Db 3120 CAGCCCAACGACTCAGCATCGACATCTACAGCGGACACTGTGTTTGGACGTTGGAGGCC 3179
Qy 3184 ACCAATACCATCAACGCTGAGCGGGAAGCCATGAGGCTGCTGGTGGTGGTGGTGGTGG 3243
Db 3180 ACCAATACCATCAACGCTGAGCGGGAAGCCATGAGGCTGCTGGTGGTGGTGGTGGTGG 3239
Qy 3244 GACCGCACAAAGCCAGGGCCATCGTCTCAACCGGACGAGGAGTACCTGTACTTCAAC 3303
Db 3240 GACCGCACAAAGCCAGGGCCATCGTCTCAACCGGACGAGGAGTACCTGTACTTCAAC 3299
Qy 3304 AACATGAGGACCGGGCAGCAAGATCGAAGCGGACGCTTGGACGCGCACAGCGCGAG 3363
Db 3300 AACATGAGGACCGGGCAGCAAGATCGAAGCGGACGCTTGGACGCGCACAGCGCGAG 3359
Qy 3364 GTCTCTTTCAACCGCGCTCATCGGCTTGTGGCTTGTGGTGGTGGTGGTGGTGGTGG 3423
Db 3360 GTCTCTTTCAACCGCGCTCATCGGCTTGTGGCTTGTGGTGGTGGTGGTGGTGGTGG 3419
Qy 3424 AAGCTGTTCTGGGTGGACGCGGACCTGAAGCGCATTTGAGAGCTGTGACCTGTCAAGGGGCC 3483
Db 3420 AAGCTGTTCTGGGTGGACGCGGACCTGAAGCGCATTTGAGAGCTGTGACCTGTCAAGGGGCC 3479
Qy 3484 AACCGCTGACCTTGGAGGACGCGCAACATCGTGCAGCTCTTGGGCTTGAACCATCTTGGC 3543
Db 3480 AACCGCTGACCTTGGAGGACGCGCAACATCGTGCAGCTCTTGGGCTTGAACCATCTTGGC 3539
Qy 3544 AAGCATCTTACTGATTCGACCGCAGCAGCAGATGATCGAGCGTGTGGAGAGAACCAAC 3603
Db 3540 AAGCATCTTACTGATTCGACCGCAGCAGCAGATGATCGAGCGTGTGGAGAGAACCAAC 3599
Qy 3604 GGGGACAAAGCGGACTCGCATCCAGGCGCTGTGGCCCACTCTGAGCATCCATGAGTGGTGG 3663
Db 3600 GGGGACAAAGCGGACTCGCATCCAGGCGCTGTGGCCCACTCTGAGCATCCATGAGTGGTGG 3659
Qy 3664 GAGGAAGTTCAGCTTGGAGGATTTCTCAGCCCAACCATGTGCTGAGCAATGTTGGTGGTGG 3723
Db 3660 GAGGAAGTTCAGCTTGGAGGATTTCTCAGCCCAACCATGTGCTGAGCAATGTTGGTGGTGG 3719
Qy 3724 TCCCAACATCTGATTTGCAAGGGTGTGGGACACCAAGTGTCTCATGCCCCAGTCCACCTC 3783
Db 3720 TCCCAACATCTGATTTGCAAGGGTGTGGGACACCAAGTGTCTCATGCCCCAGTCCACCTC 3779

```
QY 3784 GTGCTCTCTGACAACTGCTGACCTGTGGAGAGCGCCACCTGCTCCCGGACCACTTT 3843
Db |||||
QY 3780 GTGCTCTCTGACAACTGCTGACCTGTGGAGAGCGCCACCTGCTCCCGGACCACTTT 3839
Db |||||
QY 3844 GCATGTGCCACAGGGAGATCGACTGTATCCCGGGGCTTGCGCTGTGACGGCTTTCC 3903
Db |||||
QY 3840 GCATGTGCCACAGGGAGATCGACTGTATCCCGGGGCTTGCGCTGTGACGGCTTTCC 3899
Db |||||
QY 3904 GAGTGCATGACAGAGGACAGAGAGGCTGCCCGGTGTGCTCCCGCCCGCAGTTC 3963
Db |||||
QY 3900 GAGTGCATGACAGAGGACAGAGAGGCTGCCCGGTGTGCTCCCGCCCGCAGTTC 3959
Db |||||
QY 3964 TCGCGCGGGGTCACTGTGTGACCTGCGCTGCGCTGCGAGCGGAGGACAGTCTCAG 4023
Db |||||
QY 3960 TCGCGCGGGGTCACTGTGTGACCTGCGCTGCGCTGCGAGCGGAGGACAGTCTCAG 4019
Db |||||
QY 4024 GACCGCTCAGAGCGGCGGACTGTGACGCCATCTGCTGCCCAACAGTTCGGGTGCG 4083
Db |||||
QY 4020 GACCGCTCAGAGCGGCGGACTGTGACGCCATCTGCTGCCCAACAGTTCGGGTGCG 4079
Db |||||
QY 4084 AGCGGCGAGTGTCTCATCAAACAGAGTGTGACTCTCTCTGTCATGGGTGTCTAT 4263
Db |||||
QY 4080 AGCGGCGAGTGTCTCT-ATCAAACAGAGTGTGACTCTCTCTGTCATGGGTGTCTAT 4258
Db |||||
QY 4144 TCGAGGAGCTCATGTGTGAATCAACAGCGGCTCAGAGCAGCGCGGCCACAGC 4203
Db |||||
QY 4139 TCGAGGAGCTCATGTGTGAATCAACAGCGGCTCAGAGCAGCGCGGCCACAGC 4198
Db |||||
QY 4204 AGTGCATCGGCGGCTCATTTGGCATCATCTCTCTCTCTGTCATGGGTGTCTAT 4263
Db |||||
QY 4199 AGTGCATCGGCGGCTCATTTGGCATCATCTCTCTCTCTGTCATGGGTGTCTAT 4258
Db |||||
QY 4264 TTTGTGTCCAGCGGTGTGTGCCAGCTGTATGCGGGGCCAAAGCGGCTTCCCGCAC 4323
Db |||||
QY 4259 TTTGTGTCCAGCGGTGTGTGCCAGCTGTATGCGGGGCCAAAGCGGCTTCCCGCAC 4318
Db |||||
QY 4324 GAGTATGTCAGCGGACCCCGACGCTGCCCTCAATTTATAGCCCGGGGCTTCCAG 4383
Db |||||
QY 4319 GAGTATGTCAGCGGACCCCGACGCTGCCCTCAATTTATAGCCCGGGGCTTCCAG 4378
Db |||||
QY 4384 CATGGCCCTTTCACAGGATCGCATGCGGAAAGTCCATGATGAGTCCGTGAGCTGATG 4443
Db |||||
QY 4379 CATGGCCCTTTCACAGGATCGCATGCGGAAAGTCCATGATGAGTCCGTGAGCTGATG 4438
Db |||||
QY 4444 GGGGCGGGGGGGTGCCTGTATGACCGGAACAGCTGACAGGGGCTTGTCCAGC 4503
Db |||||
QY 4439 GGGGCGGGGGGGTGCCTGTATGACCGGAACAGCTGACAGGGGCTTGTCCAGC 4498
Db |||||
QY 4504 AGCTGTCCAGCAAGGACACGCTGTACCGCGATCTGAAACCGCGCTTCCCGC 4563
Db |||||
QY 4499 AGCTGTCCAGCAAGGACACGCTGTACCGCGATCTGAAACCGCGCTTCCCGC 4558
Db |||||
QY 4564 GCCACGACCCCTCTCTGTACAAATGACATGTTCTACTCTTAAACATTCGCGCACT 4623
Db |||||
QY 4559 GCCACGACCCCTCTCTGTACAAATGACATGTTCTACTCTTAAACATTCGCGCACT 4618
Db |||||
QY 4624 GCGAGACGCTACAGCCCTTACATCATTTAGGAATGGGCGCCCGCCAGACGCGCTGACG 4683
Db |||||
QY 4619 GTGAGACGCTACAGCCCTTACATCATTTAGGAATGGGCGCCCGCCAGACGCGCTGACG 4678
Db |||||
QY 4684 ACCGACGCTGTGACAGGACCTACAGCGCCAGCGCTGGAAGGCCAGCAAGTACTACCTG 4743
Db |||||
QY 4679 ACCGACGCTGTGACAGGACCTACAGCGCCAGCGCTGGAAGGCCAGCAAGTACTACCTG 4738
Db |||||
QY 4744 GATTTGAACTCGGACTACAGCCCTTATCCACCCCAACCGCCCAAGCCAGTACCTG 4803
Db |||||
QY 4739 GATTTGAACTCGGACTACAGCCCTTATCCACCCCAACCGCCCAAGCCAGTACCTG 4798
Db |||||
QY 4804 TCGGCGGAGGACGTGCGGCGCTTCCCGCCAGCGGAGGAGTACTTCCATCTCTTC 4863
Db |||||
QY 4799 TCGGCGGAGGACGTGCGGCGCTTCCCGCCAGCGGAGGAGTACTTCCATCTCTTC 4858
Db |||||
QY 4864 CCGCCCTCTCGCTCCCTGCAACGAGTCTATCTGACCTCGGCGGCGCACTCTGCTTC 4923
Db |||||
```

```
Db 4859 CCGCCCTCTCGTCCCTGACGAGCTCATCTCTGACCTCGCGGCGCACTCTGGCTTC 4918
QY 4924 TCTGTGCCCTGTAAATAGTTTTAAATATGAAACAAAGAAAAAATATATTTATGATTTA 4983
Db |||||
QY 4919 TCTGTGCCCTGTAAATAGTTTTAAATATGAAACAAAGAAAAAATATATTTATGATTTA 4978
Db |||||
QY 4984 AAAATAAATAATATGGGATTTTAAAAAATGAGAAATGTGAATGTGATGGGTGGGC 5043
Db |||||
QY 4979 AAAATAAATAATATGGGATTTTAAAAAATGAGAAATGTGAATGTGATGGGTGGGC 5038
Db |||||
QY 5044 AGGCTGCGGAGAACTTTGTA 5063
Db |||||
QY 5039 AGGCTGCGGAGAACTTTGTA 5058
Db |||||

RESULT 13
AR309179 AR309179 5098 bp DNA linear PAT 12-JUN-2003
LOCUS Sequence 1 from patent US 6555654.
DEFINITION AR309179
ACCESSION AR309179
VERSION AR309179.1 GI:31701184
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 5098)
AUTHORS Todd, J.A., Hess, J.W., Caskey, C.T., Cox, R.D., Gerhold, D.,
Hammond, H., Hey, P., Kawaguchi, Y., Merriman, T.R., Metzker, M.L.,
Nakagawa, Y., Phillips, M.S. and Twells, R.C.J.
TITLE LDL-receptor
JOURNAL Patent: US 6555654-A 1 29-APR-2003;
FEATURES Location/Qualifiers
source 1..5098
/organism="unknown"
/mol_type="genomic DNA"

Query Match 99.3%; Score 5028; DB 6; Length 5098;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 5053; Conservative 0; Mismatches 5; Indels 2; Gaps 2;

QY 4 ATGAGAGCCCGAGTGAGCGCGCGCGCCGCTCCGCGCGCGGACAAACATGAGGACGCG 63
Db 1 ATGAGAGCCCGAGTGAGCGCGCGCGCCGCTCCGCGCGCGGACAAACATGAGGACGCG 60
QY 64 CCGCCCGCGCGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 123
Db 61 CCGCCCGCGCGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
QY 124 TGCCCGCGCGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 183
Db 121 TGCCCGCGCGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
QY 184 CTGCTGAGACCGCGCGCGAGTCAAGCTGGAGTCCACCATCGTGGTTCAGCGGCTGGAGAT 243
Db 181 CTGCTGAGACCGCGCGAGTCAAGCTGGAGTCCACCATCGTGGTTCAGCGGCTGGAGAT 240
QY 244 CGCGCCGCGAGTGGACTTCCAGTTTTCAGAGGAGCCGCTGTAATCGACAGACGTGAGCGAG 303
Db 241 CGCGCCGCGAGTGGACTTCCAGTTTTCAGAGGAGCCGCTGTAATCGACAGACGTGAGCGAG 300
QY 304 GAGGCCATCAGACGACCTACTGAAACAGAGCGGGCGCGCGCTGACAGAACTGGTCTATC 363
Db 301 GAGGCCATCAGACGACCTACTGAAACAGAGCGGGCGCGCGCTGACAGAACTGGTCTATC 360
QY 364 TCCGCGCTGGTCTCTCCCGAGCGCTCGCTCGACTGGGTGGGCAAGAACTGCTACTGG 423
Db 361 TCCGCGCTGGTCTCTCCCGAGCGCTCGCTCGACTGGGTGGGCAAGAACTGCTACTGG 420
QY 424 ACAGACTCAGAGACAAACCGCATCGAGGTGGCCAACTCAATGGCACATCCCGGAAGGTG 483
Db 421 ACAGACTCAGAGACAAACCGCATCGAGGTGGCCAACTCAATGGCACATCCCGGAAGGTG 480
```

QY 484 CTCCTTCTGGCAGGACCTTGAACAGCCTTAGGGCCATCGCCTTGGACCCCGCTCACGGGTAC 543
DB 481 CTCCTTCTGGCAGGACCTTGAACAGCAGGAGGCCATCGCCTTGGACCCCGCTCACGGGTAC 540
QY 544 ATGTACTGGACAGACTGGGGTGGAGACGCCCGCGATTGAGCGGGCAGGATGGATGGCAGC 603
DB 541 ATGTACTGGACAGACTGGGGTGGAGACGCCCGCGATTGAGCGGGCAGGATGGATGGCAGC 600
QY 604 ACCCGGAGAGATCATTTCTGGACTCGACATTTTACTTGGCCCAATGGACTGACCATCGACCTG 663
DB 601 ACCCGGAGAGATCATTTCTGGACTCGACATTTTACTTGGCCCAATGGACTGACCATCGACCTG 660
QY 664 GAGGAGCAGAAAGCTCTACTTGGGCTGACGCCAAGCTTCAGCTTCATCCACCGTGCACCACTG 723
DB 661 GAGGAGCAGAAAGCTCTACTTGGGCTGACGCCAAGCTTCAGCTTCATCCACCGTGCACCACTG 720
QY 724 GACGGCTCTGTTCCGGCAGAAAGTGTGTGAGGAGCGCTTGACGACCCCTTGCCTCTGACG 783
DB 721 GACGGCTCTGTTCCGGCAGAAAGTGTGTGAGGAGCGCTTGACGACCCCTTGCCTCTGACG 780
QY 784 CTCCTCGGGGACACTCTGTACTTGGACAGACTGGCAGACCCGCTCCATCCATGCTGCAAC 843
DB 781 CTCCTCGGGGACACTCTGTACTTGGACAGACTGGCAGACCCGCTCCATCCATGCTGCAAC 840
QY 844 AAGCGCACTGGGGGGAAGAGGAGAGATCTCTGAGTGCCCTCTACTCACCCATGGACATC 903
DB 841 AAGCGCACTGGGGGGAAGAGGAGAGATCTCTGAGTGCCCTCTACTCACCCATGGACATC 900
QY 904 CAGGTGTGAGCAGGAGCGGAGCGCTTTCTTTCACAATCTGCTGTGAGGAGCAATGGC 963
DB 901 CAGGTGTGAGCAGGAGCGGAGCGCTTTCTTTCACAATCTGCTGTGAGGAGCAATGGC 960
QY 964 GGCTGCTCCACCTGTGCTGCTGCTCCCAAGCGAGCGCTTTCTACATGCGCTGCCCC 1023
DB 961 GGCTGCTCCACCTGTGCTGCTGCTCCCAAGCGAGCGCTTTCTACATGCGCTGCCCC 1020
QY 1024 ACGGCTGTGACAGTGCAGACAAACGGCAGGAGCTGTAAAGCAGGAGCGCGAGGAGTGCTG 1083
DB 1021 ACGGCTGTGACAGTGCAGACAAACGGCAGGAGCTGTAAAGCAGGAGCGCGAGGAGTGCTG 1080
QY 1084 CTGCTGGCCCGGCGAGCGACCTTACGAGAGATCTCGCTTGGACACGCCCGGACTTACCGAC 1143
DB 1081 CTGCTGGCCCGGCGAGCGACCTTACGAGAGATCTCGCTTGGACACGCCCGGACTTACCGAC 1140
QY 1144 ATCGTGTGACAGTGCAGACATCCGACCGCATTTGCCATCGACTACGACCGCTAGAG 1203
DB 1141 ATCGTGTGACAGTGCAGACATCCGACCGCATTTGCCATCGACTACGACCGCTAGAG 1200
QY 1204 GGCTATGTCTACTGACACAGATGACGAGGTGCGGGCCATCCGACGGGCGTACTTGGACGGG 1263
DB 1201 GGCTATGTCTACTGACACAGATGACGAGGTGCGGGCCATCCGACGGGCGTACTTGGACGGG 1260
QY 1264 TCTGGGGCGCAGACGCTGGTCAACACCGAGATCAACGACCCCGATGGCATCGCGTGCAC 1323
DB 1261 TCTGGGGCGCAGACGCTGGTCAACACCGAGATCAACGACCCCGATGGCATCGCGTGCAC 1320
QY 1324 TGGGTGGCCGCAAACTCTACTGGACCGACACCGGCAAGGACCGCATCGAGGTGACGGC 1383
DB 1321 TGGGTGGCCGCAAACTCTACTGGACCGACACCGGCAAGGACCGCATCGAGGTGACGGC 1380
QY 1384 CTCACAGGCACCTCCCGCAAGATCTGTGTGCGAGGACCTGGACAGCCCGGACCATC 1443
DB 1381 CTCACAGGCACCTCCCGCAAGATCTGTGTGCGAGGACCTGGACAGCCCGGACCATC 1440
QY 1444 GCATGCAACCCCGTATGGGCTCATGTACTGGACAGACTGGGGGAGAAACCTTAAATC 1503
DB 1441 GCATGCAACCCCGTATGGGCTCATGTACTGGACAGACTGGGGGAGAAACCTTAAATC 1500
QY 1504 GAGTGTGCCAATCTGGATGGGCGAGAGCGGCTGTGCTGGTCAATGCTTCCCTCGGTTGG 1563
DB 1501 GAGTGTGCCAATCTGGATGGGCGAGAGCGGCTGTGCTGGTCAATGCTTCCCTCGGTTGG 1560

QY 1564 CCCAACGGCCTGGCCCTGGACCTTGACGAGGGGAGCTTACTTGGGAGAGCCCAAGACA 1623
DB 1561 CCCAACGGCCTGGCCCTGGACCTTGACGAGGGGAGCTTACTTGGGAGAGCCCAAGACA 1620
QY 1624 GACAGATCGAGGTGATCAATGTTGATGGGACGAAGAGCGGACCCCTCTCTGAGAGCAAG 1683
DB 1621 GACAGATCGAGGTGATCAATGTTGATGGGACGAAGAGCGGACCCCTCTCTGAGAGCAAG 1680
QY 1684 CTCCTCGCACATTTTTCGGGTTCAACGCTGCTGGGGACTTTCATCTACTGACTGACTGGCAG 1743
DB 1681 CTCCTCGCACATTTTTCGGGTTCAACGCTGCTGGGGACTTTCATCTACTGACTGACTGGCAG 1740
QY 1744 CGCCGACGATCGAGCGGGTGCACAAAGTCAAGGCCAGCCGCGACGTTCATCTATTCACCG 1803
DB 1741 CGCCGACGATCGAGCGGGTGCACAAAGTCAAGGCCAGCCGCGACGTTCATCTATTCACCG 1800
QY 1804 CTGCGCGACTGATGGGGCTCAAAAGCTGTGTAATGTGGCAAGGTCTGTCGGAACCAACCGG 1863
DB 1801 CTGCGCGACTGATGGGGCTCAAAAGCTGTGTAATGTGGCAAGGTCTGTCGGAACCAACCGG 1860
QY 1864 TGTGCGGACAGAAACGGGGGTGACGCCACCTGTGCTTCTTTCACACCCACGCAACCCGG 1923
DB 1861 TGTGCGGACAGAAACGGGGGTGACGCCACCTGTGCTTCTTTCACACCCACGCAACCCGG 1920
QY 1924 TGTGGCTGCCCATATCGGCTCGGAGCTGCTGAGTGAATGAAGACCTGCTGCTGCTGAG 1983
DB 1921 TGTGGCTGCCCATATCGGCTCGGAGCTGCTGAGTGAATGAAGACCTGCTGCTGCTGAG 1980
QY 1984 GCCTTCTTGGTCTTCAACAGCAGAGCCGCTCATTCACAGGATCTCTCTCGAGACCAATAC 2043
DB 1981 GCCTTCTTGGTCTTCAACAGCAGAGCCGCTCATTCACAGGATCTCTCTCGAGACCAATAC 2040
QY 2044 AACGAGCTGGCCATCCGCTCACGGGGTCAAGGAGGCTCAGCCCTCGACCTTTCATGTC 2103
DB 2041 AACGAGCTGGCCATCCGCTCACGGGGTCAAGGAGGCTCAGCCCTCGACCTTTCATGTC 2099
QY 2104 TCCAAACCAACATCTACTGACACAGAGCTCAGCCCTGAAGACCATCAGCCGCGCTTCATG 2163
DB 2100 TCCAAACCAACATCTACTGACACAGAGCTCAGCCCTGAAGACCATCAGCCGCGCTTCATG 2159
QY 2164 AACGGGAGCTCGGTGGAGCAGCTGGTGGAGTTTGGCTTGTACTACCCCGAGGGCATGGCC 2223
DB 2160 AACGGGAGCTCGGTGGAGCAGCTGGTGGAGTTTGGCTTGTACTACCCCGAGGGCATGGCC 2219
QY 2224 GTTGACTGAGTGGCAGAAACCTCTACTGGGCCGACACTGGGACCAACAGCAATCGAAGTG 2283
DB 2220 GTTGACTGAGTGGCAGAAACCTCTACTGGGCCGACACTGGGACCAACAGCAATCGAAGTG 2279
QY 2284 CGCGGCTGGACGGGCGAGTTCCGGCAAGTCTCGTGTGGAGGAGCTTGGACAAACCCGAGG 2343
DB 2280 CGCGGCTGGACGGGCGAGTTCCGGCAAGTCTCGTGTGGAGGAGCTTGGACAAACCCGAGG 2339
QY 2344 TCGCTGGCCCTGGATCCCAACGAGGCTACATCTACTGGAACGAGTGGGGGCGCAAGCGG 2403
DB 2340 TCGCTGGCCCTGGATCCCAACGAGGCTACATCTACTGGAACGAGTGGGGGCGCAAGCGG 2399
QY 2404 AGGATCGTGGGGCTTCATGGAACGGGACCAACTGCTGATGACGCTGGTGGACAAAGTTGGGC 2463
DB 2400 AGGATCGTGGGGCTTCATGGAACGGGACCAACTGCTGATGACGCTGGTGGACAAAGTTGGGC 2459
QY 2464 CGGGCCAAACGACCTCACCATTTGACTACGCTGACCGAGCCCTCTACTGGACCGACCTGGAC 2523
DB 2460 CGGGCCAAACGACCTCACCATTTGACTACGCTGACCGAGCGCCCTCTACTGGACCGACCTGGAC 2519
QY 2524 ACCAAACATGATCGAGTGTGTCACATGCTGGGTGACGAGCGGGTCTGATTTGCCACGAT 2583
DB 2520 ACCAAACATGATCGAGTGTGTCACATGCTGGGTGACGAGCGGGTCTGATTTGCCACGAT 2579
QY 2584 CTCCTCGCACCCGTTCTGGTCTGACGAGTACAGGATTTATATCTACTGACAGACTGGGAAT 2643
DB 2580 CTCCTCGCACCCGTTCTGGTCTGACGAGTACAGGATTTATATCTACTGACAGACTGGGAAT 2639
QY 2644 CTGACACAGCAATTGAGCGGGCGGACAAAGACTAGCGGCGGAAACCGCACCCCTCATTCAGGGC 2703

||||| 2640 CTGCAAGATTTAGCGGGCCGACAGACTAGCGGCGGAAACGACCCCTCATCCAGGGC 2699
QY ||||| 2704 CACCTGGACTTCGTGATGACATCTCTGGTGTCCACTCTCCCGCAGAGATGGCTCAAT 2763
Db ||||| 2700 CACCTGGACTTCGTGATGACATCTCTGGTGTCCACTCTCCCGCAGAGATGGCTCAAT 2759
QY ||||| 2764 GACTGTATGCACAAACAGGCGAGTGTGGCGAGCTGTGGCTTGGCCATCCCGCGGGCCAC 2823
Db ||||| 2760 GACTGTATGCACAAACAGGCGAGTGTGGCGAGCTGTGGCTTGGCCATCCCGCGGGCCAC 2819
QY ||||| 2824 CGCTGGGTGGGCTTCACACTACACCTCTGACCCCGCAGCGCGCAACTGACAGCCGGCC 2883
Db ||||| 2820 CGCTGGGTGGGCTTCACACTACACCTCTGACCCCGCAGCGCGCAACTGACAGCCGGCC 2879
QY ||||| 2884 AACACTTCTTCTGTGTAGCAGCAAAATCTGCCATCAGTCCGATGATCCCGACACAG 2943
Db ||||| 2880 AACACTTCTTGTGTAGCAGCAAAATCTGCCATCAGTCCGATGATCCCGACACAG 2939
QY ||||| 2944 CACAGCCGGATCTCATCTCTGCCCTCTGCATGACCTGAGGAAGCTCAAGCCATCAGTAT 3003
Db ||||| 2940 CACAGCCGGATCTCATCTCTGCCCTCTGCATGACCTGAGGAAGCTCAAGCCATCAGTAT 2999
QY ||||| 3004 GACCACTTGGACAAAGTTTCATCTACTGGGTGGATGGCGCCAGAAATCAAGCGAGCCAA 3063
Db ||||| 3000 GACCACTTGGACAAAGTTTCATCTACTGGGTGGATGGCGCCAGAAATCAAGCGAGCCAA 3059
QY ||||| 3064 GAGAGCGGACCCAGGCTTTGTTTGGTGTAGCTCTCTGAGCAAGGCAAAACCCAGACAG 3123
Db ||||| 3060 GAGAGCGGACCCAGGCTTTGTTTGGTGTAGCTCTCTGAGCAAGGCAAAACCCAGACAG 3119
QY ||||| 3124 CAGCCCCAGACTCAGCATTCAGATCTACAGCGCGGACACTGTTCTGACGTCGAGGCC 3183
Db ||||| 3120 CAGCCCCAGACTCAGCATTCAGATCTACAGCGCGGACACTGTTCTGACGTCGAGGCC 3179
QY ||||| 3184 ACCAATACCATCAACGTTCCACAGGCTGAGCGGGGAAGCATGGGGGTGGTCTGGTGG 3243
Db ||||| 3180 ACCAATACCATCAACGTTCCACAGGCTGAGCGGGGAAGCATGGGGGTGGTCTGGTGG 3239
QY ||||| 3244 GACCGGACAGCCAGGCGCATCTGTCTGAACGCGAGCGAGGTTACTGTACTTACC 3303
Db ||||| 3240 GACCGGACAGCCAGGCGCATCTGTCTGAACGCGAGCGAGGTTACTGTACTTACC 3299
QY ||||| 3304 AACATGAGGACCGGCGAGCAAGATCGAAACGCGAGCCCTGAGCGGACCGAGCGCGAG 3363
Db ||||| 3300 AACATGAGGACCGGCGAGCAAGATCGAAACGCGAGCCCTGAGCGGACCGAGCGCGAG 3359
QY ||||| 3364 GTCTCTTTCACACCGGCTCATCTCGCCCTGTGGCCCTGGTGGTAGAACACACACTGGGC 3423
Db ||||| 3360 GTCTCTTTCACACCGGCTCATCTCGCCCTGTGGCCCTGGTGGTAGAACACACACTGGGC 3419
QY ||||| 3424 AAGCTGTTCTGGGTGGAACGCACTGAAGCGCATGAGAGCTGTGACTGTGAGGGGCC 3483
Db ||||| 3420 AAGCTGTTCTGGGTGGAACGCACTGAAGCGCATGAGAGCTGTGAGCTGTGAGGGGCC 3479
QY ||||| 3484 AACCCCTGACCTTGAGAGCGCAACATCTGTGAGCTCTGGGCTGACCATCTCTGGC 3543
Db ||||| 3480 AACCCCTGACCTTGAGAGCGCAACATCTGTGAGCTCTGGGCTGACCATCTCTGGC 3539
QY ||||| 3544 AAGCATCTTACTGATCGACCGCAGCAGCAGATGATCGAGCGTGTGGAGAGACCAACC 3603
Db ||||| 3540 AAGCATCTTACTGATCGACCGCAGCAGCAGATGATCGAGCGTGTGGAGAGACCAACC 3599
QY ||||| 3604 GGGGACAGCGGACTCGCATCCAGGGCGGTGCGCCACCTCACTGGCATCCATGCACTG 3663
Db ||||| 3600 GGGGACAGCGGACTCGCATCCAGGGCGGTGCGCCACCTCACTGGCATCCATGCACTG 3659
QY ||||| 3664 GAGGAAGTCAGCTTGAGAGGTTCTCAGCCACCATGTGCCCGTGACATGTTGGCTGC 3723
Db ||||| 3660 GAGGAAGTCAGCTTGAGAGGTTCTCAGCCACCATGTGCCCGTGACATGTTGGCTGC 3719
QY ||||| 3724 TCCCAATCTGTATTGCCAAGGATGGGACACCAACGCTGCTCATGCCCATGTCACCTC 3783

Db ||||| 3720 TCCCAATCTGTATTGCCAAGGATGAGGACACCAACGCTGCTCATGCCAGTCCACCTC 3779
QY ||||| 3784 GTGCTCTTGCAAAACCTGCTGACCTGTGAGAGAGCGGCCCACTGCTCCCGGACAGTATT 3843
Db ||||| 3780 GTGCTCTTGCAAAACCTGCTGACCTGTGAGAGAGCGGCCACCTGCTCCCGGACAGTATT 3839
QY ||||| 3844 GCATGTGCCACAGGGGAGATCGACTGTATCCCGGGGCTCGCGCTGTGACGCGCTTTCCC 3903
Db ||||| 3840 GCATGTGCCACAGGGGAGATCGACTGTATCCCGGGGCTCGCGCTGTGACGCGCTTTCCC 3899
QY ||||| 3904 GAGTGCATGATCAGAGAGCAGAGAGGCTGCCCGCTGTGCTCCCGCGCCAGTTCGCC 3963
Db ||||| 3900 GAGTGCATGATCAGAGAGCAGAGAGGCTGCCCGCTGTGCTCCCGCGCCAGTTCGCC 3959
QY ||||| 3964 TCGCGCGGGGTCAAGTGTGTGGAACCTGCGCTTCGCGTCGACGCGAGGAGACTGTG 4023
Db ||||| 3960 TCGCGCGGGGTCAAGTGTGTGGAACCTGCGCTTCGCGTCGACGCGAGGAGACTGTG 4019
QY ||||| 4024 GACCGCTCAGACGAGGGGACTGTGACGCACTCTGCTGCCCAACCACTTCGGGTGCGG 4083
Db ||||| 4020 GACCGCTCAGACGAGGGGACTGTGACGCACTCTGCTGCCCAACCACTTCGGGTGCGG 4079
QY ||||| 4084 AGCGGCCAGTGTGCTCTCATCAAAACAGCAGTGCAGTCTCTTCCCGCACTGTATCGACG 4143
Db ||||| 4080 AGCGGCCAGTGTGCTCTCATCAAAACAGCAGTGCAGTCTCTTCCCGCACTGTATCGACG 4138
QY ||||| 4144 TCGGACGAGCTCATGTGTGAAATCAACAGCGGCTCTCAGACGAGCGCGGCCACAGC 4203
Db ||||| 4139 TCGGACGAGCTCATGTGTGAAATCAACAGCGGCTCTCAGACGAGCGCGGCCACAGC 4198
QY ||||| 4204 AGTGCATCTGGGCGGCTCATTTGGCATCATCTCTCTCTTCTGTCATGGGTGCTCTAT 4263
Db ||||| 4199 AGTGCATCTGGGCGGCTCATTTGGCATCATCTCTCTCTTCTGTCATGGGTGCTCTAT 4259
QY ||||| 4264 TTTGTGTGCCAGCGTGTGTCAGCGCTATGCGGGGCAACAGGGCCCTTCCCGCAC 4323
Db ||||| 4259 TTTGTGTGCCAGCGTGTGTCAGCGCTATGCGGGGCAACAGGGCCCTTCCCGCAC 4318
QY ||||| 4324 GAGTATGTGACGGGACCCCGCAGCTGCTGCTCAATTTCTATAGCCCGGGCGGTTCGAC 4383
Db ||||| 4319 GAGTATGTGACGGGACCCCGCAGCTGCTGCTCAATTTCTATAGCCCGGGCGGTTCGAC 4378
QY ||||| 4384 CATGGCCCTTTCACAGGATCGCATGCGGAAAGTCTATGATGCTCCGTGAGCTGATG 4443
Db ||||| 4379 CATGGCCCTTTCACAGGATCGCATGCGGAAAGTCTATGATGAGCTCCGTGAGCTGATG 4438
QY ||||| 4444 GGGGCGGGGCGGGGTGCCCTGTACGACGGAACCACTGTCACAGGGGCTCTGTTCACG 4503
Db ||||| 4439 GGGGCGGGGCGGGGTGCCCTGTACGACGGAACCACTGTCACAGGGGCTCTGTTCACG 4498
QY ||||| 4504 AGCTGTCCAGCAGGAAGGCCACGCTGTACCGCGGATCTCTGAACCGCGGCTTCGCG 4563
Db ||||| 4499 AGCTGTCCAGCAGGAAGGCCACGCTGTACCGCGGATCTCTGAACCGCGGCTTCGCG 4558
QY ||||| 4564 GCCAGGACCCCTCCCTGTACAAATGGAATGTTCTACTCTTCAAAATTCGCGGCACT 4623
Db ||||| 4559 GCCAGGACCCCTCCCTGTACAAATGGAATGTTCTACTCTTCAAAATTCGCGGCACT 4618
QY ||||| 4624 GCGAGACGTTACAGGCTTACATCATTCGAGGAATGGCGCCCGCAGACGCGCTTCAGC 4683
Db ||||| 4619 GTGAGACGTTACAGGCTTACATCATTCGAGGAATGGCGCCCGCAGACGCGCTTCAGC 4678
QY ||||| 4684 ACCGACGTGTGACAGGACTACAGCGCGGATCTGTAACCGCGGATCTCTGAACCGCGGCTTCGCG 4743
Db ||||| 4679 ACCGACGTGTGACAGGACTACAGCGCGGATCTGTAACCGCGGATCTCTGAACCGCGGCTTCGCG 4738
QY ||||| 4744 GATTTGAATCGGACTCAGACCCCTATCCACCCCGCAGCGCCCGCAGCAGTACCTG 4803
Db ||||| 4739 GATTTGAATCGGACTCAGACCCCTATCCACCCCGCAGCGCCCGCAGCAGTACCTG 4798
QY ||||| 4804 TCGCGGAGGACAGCTGCGCGGCTCTGCGCGGCTCCAGAGAGGAGTACTTCCATCTCTTC 4863
Db ||||| 4799 TCGCGGAGGACAGCTGCGCGGCTCTGCGCGGCTCCAGAGAGGAGTACTTCCATCTCTTC 4858

Db	1261	TCGCGGCGCAGACGCTGTCTCAACCCGAGATCAACGACCCCGATCGGCTCGAC	1320
QY	1324	TGGGTGGCCGGAACCTCTTACTGGACCGGACACGCGGACCGGACCGGATCGAGGTGAGCGC	1383
Db	1321	TGGGTGGCCGGAACCTCTTACTGGACCGGACACGCGGACCGGATCGAGGTGAGCGC	1380
QY	1384	CTCAACGGCACCTCCCGCAAGATCTCTGTGTGTGGAGGACCTGGAGAGCCCGAGGCCATC	1443
Db	1381	CTCAACGGCACCTCCCGCAAGATCTCTGTGTGTGGAGGACCTGGAGAGCCCGAGGCCATC	1440
QY	1444	GCACTGCAACCCCGTGTGATGGGCTCTATGTACTGGACAGACTGGGGAGAGAACCTTAAATC	1503
Db	1441	GCACTGCAACCCCGTGTGATGGGCTCTATGTACTGGACAGACTGGGGAGAGAACCTTAAATC	1500
QY	1504	GAGTGTGCAACTTGTGATGGGACGAGAGCGGCTGTGTGTGTGGAGGACCTGGAGAGCCCGAGGCCATC	1563
Db	1501	GAGTGTGCAACTTGTGATGGGACGAGAGCGGCTGTGTGTGTGGAGGACCTGGAGAGCCCGAGGCCATC	1560
QY	1564	CCCAACGGGCTGGCCCTGTGACCTGACGAGGAGGAAAGCTCTACTGGGGAGAGCCCGAGGACA	1623
Db	1561	CCCAACGGGCTGGCCCTGTGACCTGACGAGGAGGAAAGCTCTACTGGGGAGAGCCCGAGGACA	1620
QY	1624	GACAGATCGAGGTGATCAATTTGATGGGACGAGAGCGGACCTCTCTGGAGGACCAAG	1683
Db	1621	GACAGATCGAGGTGATCAATTTGATGGGACGAGAGCGGACCTCTCTGGAGGACCAAG	1680
QY	1684	CTCCCGCACATTTTCGGGTTCACGCTGTGGGGGACTTCACTACTGACTGACTGGGAG	1743
Db	1681	CTCCCGCACATTTTCGGGTTCACGCTGTGGGGGACTTCACTACTGACTGACTGGGAG	1740
QY	1744	CGCCGACGATCGAGCGGGTGCAAGGTCAGAGGCGGAGAGCTCATCATTCACACAG	1803
Db	1741	CGCCGACGATCGAGCGGGTGCAAGGTCAGAGGCGGAGAGCTCATCATTCACACAG	1800
QY	1804	CTGCCCGACCTGATGGGGCTCAAGCTGTGAATGTGGCCAGGTCGTGGAAACCAACCG	1863
Db	1801	CTGCCCGACCTGATGGGGCTCAAGCTGTGAATGTGGCCAGGTCGTGGAAACCAACCG	1860
QY	1864	TGTGGGACAGAAACGGGGGTGACGACCTGTGCTTCTTACACCCCGACGACCCGG	1923
Db	1861	TGTGGGACAGAAACGGGGGTGACGACCTGTGCTTCTTACACCCCGACGACCCGG	1920
QY	1924	TGTGGCTGCCCATCGGCTGGAGCTGTGAGTGACATGAAGACCTGCTGCTGCTGAG	1983
Db	1921	TGTGGCTGCCCATCGGCTGGAGCTGTGAGTGACATGAAGACCTGCTGCTGCTGAG	1980
QY	1984	GCCTTCTTGGTCTTACAGAGAGCGGCTATCCAGAGATCTCCCTCGAGACCAATTAAC	2043
Db	1981	GCCTTCTTGGTCTTACAGAGAGCGGCTATCCAGAGATCTCCCTCGAGACCAATTAAC	2040
QY	2044	AACGACGTGGCCATCCGCTCACGGGCGTCAAGGAGGCTCAGCCCTGGAATTTGATGTG	2103
Db	2041	AACGACGTGGCCATCCGCTCACGGGCGTCAAGGAGGCTCAGCCCTGGAATTTGATGTG	2099
QY	2104	TCCAAACACCATCTACTGGAACAGCTGAGTGAAGACCATCAGCGGCGCTTCAATG	2163
Db	2100	TCCAAACACCATCTACTGGAACAGCTGAGTGAAGACCATCAGCGGCGCTTCAATG	2159
QY	2164	AACGGGAGCTCGTGGAGGACGTTGTGAGTTTGGCTTGAATACCCGAGGCGATGGCC	2223
Db	2160	AACGGGAGCTCGTGGAGGACGTTGTGAGTTTGGCTTGAATACCCGAGGCGATGGCC	2219
QY	2224	GTTGACTGGATGGGCAAGAACTCTACTGGGCGGACACTGGGACCAACAGAAATCGAAGTG	2283
Db	2220	GTTGACTGGATGGGCAAGAACTCTACTGGGCGGACACTGGGACCAACAGAAATCGAAGTG	2279
QY	2284	GCGGGCTGGAACGGGAGTTCCGGCAAGTCTCTGTGTGGAGGAGCTTGGACAAACCCGAGG	2343
Db	2280	GCGGGCTGGAACGGGAGTTCCGGCAAGTCTCTGTGTGGAGGAGCTTGGACAAACCCGAGG	2339
QY	2344	TCGCTGGCCCTGGATCCCAACGGGCTTACTTACTGGACCGAGTGGGGCGGACCGG	2403
Db	2340	TCGCTGGCCCTGGATCCCAACGGGCTTACTTACTGGACCGAGTGGGGCGGACCGG	2399
QY	2404	AGGATCGTGGCGGCTTTCATGGAACGGGACCAACTGATGACGCTGGTGGACAAAGGTGGGC	2463
Db	2400	AGGATCGTGGCGGCTTTCATGGAACGGGACCAACTGATGACGCTGGTGGACAAAGGTGGGC	2459
QY	2464	CGGGCCAAAGACCTCACCATTTGACTGACCTGACGAGCGCTTACTTGGACCGACCTGGAC	2523
Db	2460	CGGGCCAAAGACCTCACCATTTGACTGACCTGACGAGCGCTTACTTGGACCGACCTGGAC	2519
QY	2524	ACCAACATGATCGAGTCTGCTCAACATGCTGGGTGACGAGCGGCTGATGATTCGCCACCAT	2583
Db	2520	ACCAACATGATCGAGTCTGCTCAACATGCTGGGTGACGAGCGGCTGATGATTCGCCACCAT	2579
QY	2584	CTCCCGGACCCGTTCTGCTGCTGACGAGTACAGCGATATATCTACTGACAGACTGGAAAT	2643
Db	2580	CTCCCGGACCCGTTCTGCTGCTGACGAGTACAGCGATATATCTACTGACAGACTGGAAAT	2639
QY	2644	CTGCACAGCATTTGAGCGGGCGGACAAAGACTAGCGGCGGAAACCGCACCTCATCCAGGCG	2703
Db	2640	CTGCACAGCATTTGAGCGGGCGGACAAAGACTAGCGGCGGAAACCGCACCTCATCCAGGCG	2699
QY	2704	CACCTGGACTTTCGTGATGGACATCTCTGTGTTCACCTCTCCCGCAGGATGGCTCAAT	2763
Db	2700	CACCTGGACTTTCGTGATGGACATCTCTGTGTTCACCTCTCCCGCAGGATGGCTCAAT	2759
QY	2764	GACTGTATGCAACAAACGGGCGAGTGGGGAGCTGTGCTTGCATCCCGCGGCGCCAC	2823
Db	2760	GACTGTATGCAACAAACGGGCGAGTGGGGAGCTGTGCTTGCATCCCGCGGCGCCAC	2819
QY	2824	CGCTGGGCTCGGCTCACACTACACCTTGACCCCGACGCGCAACTGACGCGCGCC	2883
Db	2820	CGCTGGGCTCGGCTCACACTACACCTTGACCCCGACGCGCAACTGACGCGCGCC	2879
QY	2884	ACCACTTCTTGTCTTTCAGCCGAAATCTGCTCATCTGATCGGATGATCCCGGACGACAG	2943
Db	2880	ACCACTTCTTGTCTTTCAGCCGAAATCTGCTCATCTGATCGGATGATCCCGGACGACAG	2939
QY	2944	CACAGCCGGATCTCATCTGCTGCGCTGACGAGTGAAGACGTCGAAGCCATGACTAT	3003
Db	2940	CACAGCCGGATCTCATCTGCTGCGCTGACGAGTGAAGACGTCGAAGCCATGACTAT	2999
QY	3004	GACCACTGGAACAAAGTTTCATCTACTGGGTGATGGGCGGACCAAGACATCAAGCGAGCCAG	3063
Db	3000	GACCACTGGAACAAAGTTTCATCTACTGGGTGATGGGCGGACCAAGACATCAAGCGAGCCAG	3059
QY	3064	GACGACGGGACCCAGCCCTTGTGCTCTCTGAGCCAAAGGCCAAAAACCGAGACAGG	3123
Db	3060	GACGACGGGACCCAGCCCTTGTGCTCTCTGAGCCAAAGGCCAAAAACCGAGACAGG	3119
QY	3124	CAGCCCGACGACCTCAGCATCGACATCTACAGCGGACACTGTTCTTGGACGTGGAGGCC	3183
Db	3120	CAGCCCGACGACCTCAGCATCGACATCTACAGCGGACACTGTTCTTGGACGTGGAGGCC	3179
QY	3184	ACCAATACCATCAACGCTCCACAGGCTGAGCGGGGAGCCATGGGGGTGGTCTGCTGGGG	3243
Db	3180	ACCAATACCATCAACGCTCCACAGGCTGAGCGGGGAGCCATGGGGGTGGTCTGCTGGGG	3239
QY	3244	GACCGGACAAAGCCCGAGGCCATCGTGTCTAAACCGGAGCGAGGATACCTGTACTTACCC	3303
Db	3240	GACCGGACAAAGCCCGAGGCCATCGTGTCTAAACCGGAGCGAGGATACCTGTACTTACCC	3299
QY	3304	AACATGAGGACCGGGGACCAAGATCGAAACGCGACCTTGGACCGGACCGAGCGGAG	3363
Db	3300	AACATGAGGACCGGGGACCAAGATCGAAACGCGACCTTGGACCGGACCGAGCGGAG	3359
QY	3364	GTCCTTTCACCAACCGGCTCATCCGCCCTGTGGCCCTGTGGTGGGACAAACACTGGGC	3423
Db	3360	GTCCTTTCACCAACCGGCTCATCCGCCCTGTGGCCCTGTGGTGGGACAAACACTGGGC	3419
QY	3424	AAGGTGTTCTGGGTGGAGCGGACCTGAAAGGCAATGAGAGCTGTGACCTGTGAGGGGCC	3483
Db	3420	AAGGTGTTCTGGGTGGAGCGGACCTGAAAGGCAATGAGAGCTGTGACCTGTGAGGGGCC	3479

Qy	3484	AACGCGCTGACCCCTGGAGAGCGCCAA	CATCGTGTGAGCGCTCTGGGCGTGAC	CCATCTCTTGGC	3543
Db	3480	AACGCGCTGACCCCTGGAGAGCGCCAA	CATCGTGTGAGCGCTCTGGGCGCTGAC	CCATCTCTTGGC	3539
Qy	3544	AAGCATCTCTACTGATCGACCGG	CCAGCAGCAGATGATCGAGCGTGTGGAGAGAC	CCACC	3603
Db	3540	AAGCATCTCTACTGATCGACCGG	CCAGCAGCAGATGATCGAGCGTGTGGAGAGAC	CCACC	3599
Qy	3604	GGGGAACAAGCGGACTCGCATC	CAGGGCCGTGTGCCCACTCACTGGCATCTCCATG	CGAGTG	3663
Db	3600	GGGGAACAAGCGGACTCGCATC	CAGGGCCGTGTGCCCACTCACTGGCATCTCCATG	CGAGTG	3659
Qy	3664	GAGGAAGTCAGCCCTGGAGAGGTT	CTCAGGCCACCCATGTGCGCCGTGACATATG	TGGTGTC	3723
Db	3660	GAGGAAGTCAGCCCTGGAGAGGTT	CTCAGGCCACCCATGTGCGCCGTGACATATG	TGGTGTC	3719
Qy	3724	TCCCACATCTGTATTGCGAAGGGT	GTATGGGACACCAAGGTGTCTATGCCCACTCCAG	TCCACCTC	3783
Db	3720	TCCCACATCTGTATTGCGAAGGGT	GTATGGGACACCAAGGTGTCTATGCCCACTCCAG	TCCACCTC	3779
Qy	3784	GTGCTCTCTGCAAACTGCTGAC	CTGTGTGAGAGCGGCCCACTGCTCCCCCGGAC	CAGTTT	3843
Db	3780	GTGCTCTCTGCAAACTGCTGAC	CTGTGTGAGAGCGGCCCACTGCTCCCCCGGAC	CAGTTT	3839
Qy	3844	GCATGTGCCACAGGGGAGATGCA	CTGTATCCCGGGGCTGTGGCTGTGACGGCTTTCCC		3903
Db	3840	GCATGTGCCACAGGGGAGATGCA	CTGTATCCCGGGGCTGTGGCTGTGACGGCTTTCCC		3899
Qy	3904	GAGTGCATGACACAGACGACGAG	GGGGCTGCCCGTGTGCTCCGCGGCCACAGTTTCCC		3963
Db	3900	GAGTGCATGACACAGACGACGAG	GGGGCTGCCCGTGTGCTCCGCGGCCACAGTTTCCC		3959
Qy	3964	TGCGCGCGGGGTCAAGTGTGTG	GAACCTCGCGCTGCGCTCGACGCGCGGACGAGCAG	ACTGTTCAG	4023
Db	3960	TGCGCGCGGGGTCAAGTGTGTG	GAACCTCGCGCTGCGCTCGACGCGCGGACGAGCAG	ACTGTTCAG	4019
Qy	4024	GACCGCTCAGACGAGGGCGGACT	GTGTGACGCCATCTGCTGCGCCACCAAGTTTCCGGTGTGCG		4083
Db	4020	GACCGCTCAGACGAGGGCGGACT	GTGTGACGCCATCTGCTGCGCCACCAAGTTTCCGGTGTGCG		4079
Qy	4084	AGGGCCAGTGTTCCTCATCAAC	ACAGCAGTGGGACTCTTCCCGGACTGTATTCGACGGC		4143
Db	4080	AGGGCCAGTGTGTCTT-ATCAA	AACAGTGGGACTCTTCCCGGACTGTATTCGACGGC		4138
Qy	4144	TCCGACGAGCTCATGTGTGAAAT	CACCAAGCGGCCCTCAGACGACGCCCGGCCACAGC		4203
Db	4139	TCCGACGAGCTCATGTGTGAAAT	CACCAAGCGGCCCTCAGACGACGCCCGGCCACAGC		4198
Qy	4204	AGTGCCATCGGGGCCGTGAT	TGGCATCATCTCTCTCTCTTCTGTGATGGTGTGTCTAT		4263
Db	4199	AGTGCCATCGGGGCCGTGAT	TGGCATCATCTCTCTCTCTTCTGTGATGGTGTGTCTAT		4258
Qy	4264	TTTGTGTGCGAGCGGTGTGTCC	ACGCGCTATGCGGGGCGCAACGGGCCCTTCCCGCAC		4323
Db	4259	TTTGTGTGCGAGCGGTGTGTCC	ACGCGCTATGCGGGGCGCAACGGGCCCTTCCCGCAC		4318
Qy	4324	GAGTATGTACGCGGACCCCGC	CACTGCCCTCAATTTTATAGCCCCGGCGGTGCCAG		4383
Db	4319	GAGTATGTACGCGGACCCCGC	CACTGCCCTCAATTTTATAGCCCCGGCGGTGCCAG		4378
Qy	4384	CATGGCCCCCTTCAAGC	CATCGCATCGGAAAGTCCATGATGAGTCCGTGAGCCTGATG		4443
Db	4379	CATGGCCCCCTTCAAGC	CATCGCATCGGAAAGTCCATGATGAGTCCGTGAGCCTGATG		4438
Qy	4444	GGGGCGCGGGCGGGGTGCC	CTGTACGACCGGAACTCAGGGGCTCGTTCAGC		4503
Db	4439	GGGGCGCGGGCGGGGTGCC	CTGTACGACCGGAACTCAGGGGCTCGTTCAGC		4498
Qy	4504	AGCTCGTCCAGCAGGAAGG	CCAGCTGTACCGCGGATCTTGAACCGCGCGCTCCCG		4563
Db	4499	AGCTCGTCCAGCAGGAAGG	CCAGCTGTACCGCGGATCTTGAACCGCGCGCTCCCG		4558
Qy	4564	GCCAAGCAACCCCTCCCTGT	GTAAACATGATGTTCTACTCTTCAACATTCGCGGCACT		4623

[illegible]

Db 121 TGCCGGCCCCCGCGCGCTCGCGCTCTCTGCTATTGTCACACCGCGCGGACGTACGG 180
Qy 184 CTGTTGGACGCGCGCGGAGTCAAGCTGGAGTCCACCATCTGTTGGTACGCGGCTTGAGGAT 243
Db 181 CTGTTGGACGCGCGCGGAGTCAAGCTGGAGTCCACCATCTGTTGGTCAAGCGGCTTGAGGAT 240
Qy 244 GCGCGCGAGTGGACTTCAGTTTTCAGAGGAGCGGTGTTGACACAGAGTGGAGCGAG 303
Db 241 GCGCGCGAGTGGACTTCAGTTTTCAGAGGAGCGGTGTTGACACAGAGTGGAGCGAG 300
Qy 304 GAGGCGCATCAAGACACTTACCTGAACACAGACGCGGCGCGCTGTCAGAACGTGTCTATC 363
Db 301 GAGGCGCATCAAGACACTTACCTGAACACAGACGCGGCGCGCTGTCAGAACGTGTCTATC 360
Qy 364 TCCGCGCTGGTCTTCCCGACGCGCTCGCTGCGACTGGGTGGGCAAGAGTGTACTGG 423
Db 361 TCCGCGCTGGTCTTCCCGACGCGCTCGCTGCGACTGGGTGGGCAAGAGTGTACTGG 420
Qy 424 ACGGACTCAGACACCAACCGCATCGAGTGGCCAACTCAATGGGCACATCCCGGAAGGTG 483
Db 421 ACGGACTCAGACACCAACCGCATCGAGTGGCCAACTCAATGGGCACATCCCGGAAGGTG 480
Qy 484 CTCTTCTGCGAGCACTTGACAGCTTAGGGCCATCGCTTGGAACCCGCTCACCGGTTAC 543
Db 481 CTCTTCTGCGAGCACTTGACAGCGGAGGCGCATCGCTTGGAACCCGCTCACCGGTTAC 540
Qy 544 ATGTACTGACAGACTGGGGTGAGACGCGCCCGGATTTAGCGGGCAGGATGGACG 603
Db 541 ATGTACTGACAGACTGGGGTGAGACGCGCCCGGATTTAGCGGGCAGGATGGATGGACG 600
Qy 604 ACCCGAGATCATTTGTGACCTCGACATTTACTGCGCCCAATGGACTGACCATCGACCTG 663
Db 601 ACCCGAAGATCATTTGTGACCTCGACATTTACTGCGCCCAATGGACTGACCATCGACCTG 660
Qy 664 GAGGAGCAGAGCTTACTTGGCTGACGCAAGCTCAGCTTATCCACCGTGGCCAACTG 723
Db 661 GAGGAGCAGAGCTTACTTGGCTGACGCAAGCTCAGCTTATCCACCGTGGCCAACTG 720
Qy 724 GACGGTCTTCCGGCAGAGTGTGGAGGGCAGCTGACGACCGCTTCCGCTGAG 783
Db 721 GACGGTCTTCCGGCAGAGTGTGGAGGGCAGCTGACGACCGCTTCCGCTGAG 780
Qy 784 CTCTCGGGGACACTCTGTACTTGGACAGACTGGCAGACCGCTTCCATCATCGCTGCAAC 843
Db 781 CTCTCGGGGACACTCTGTACTTGGACAGACTGGCAGACCGCTTCCATCATCGCTGCAAC 840
Qy 844 AAGCGCACTGGGGGGAAGGAGATCTGTAGTGCCTTATCTACCCATGGACATC 903
Db 841 AAGCGCACTGGGGGGAAGGAGAGATCTGTAGTGCCTTATCTACCCATGGACATC 900
Qy 904 CAGGTGCTGAGCCAGGAGCGGACCTTTCTTCCACACTCTGTTGAGGAGCAATGGC 963
Db 901 CAGGTGCTGAGCCAGGAGCGGACCTTTCTTCCACACTCTGTTGAGGAGCAATGGC 960
Qy 964 GGCTGCTCCACCTGTGCTGTCTCCCAAGCGAGCTTTTACACATGCGCTGCCCC 1023
Db 961 GGCTGCTCCACCTGTGCTGTCTCCCAAGCGAGCTTTTACACATGCGCTGCCCC 1020
Qy 1024 ACGGGTGTGAGTGCAGACAAACGCGAGAGCTGTAAAGGAGGAGCGGAGAGTGTG 1083
Db 1021 ACGGGTGTGAGTGCAGACAAACGCGAGAGCTGTAAAGGAGGAGCGGAGAGTGTG 1080
Qy 1084 CTGCTGGCGCGGAGCGGACCTTACGGAGATCTCGCTGGACACGCGGACTTACCGGAC 1143
Db 1081 CTGCTGGCGCGGAGCGGACCTTACGGAGATCTCGCTGGACACGCGGACTTACCGGAC 1140
Qy 1144 ATCGTGTGAGTGGACAGACATCCGCGACCGCATTTGCCATCGACTACGACCGCTTAGAG 1203
Db 1141 ATCGTGTGAGTGGACAGACATCCGCGACCGCATTTGCCATCGACTACGACCGCTTAGAG 1200
Qy 1204 GGCTATGTCTTACTGAGACAGATGAGAGTGGGGCCATCCGAGGGGCTACTTGGACGGG 1263
Db 1201 GGCTATGTCTTACTGAGACAGATGAGAGTGGGGCCATCCGAGGGGCTACTTGGACGGG 1260

Qy 1264 TCTGGGCGCAGAGCGTGTCAACACCGAGATCAACGACCCCGATGTCATCGCGGTGCAC 1323
Db 1261 TCTGGGCGCAGAGCGTGTCAACACCGAGATCAACGACCCCGATGTCATCGCGGTGCAC 1320
Qy 1324 TGGGTGCGCCGAAACCTTACTGTGACCGACACGCGGCAACGACCGCATCGAGGTGACGCGC 1383
Db 1321 TGGGTGCGCCGAAACCTTACTGTGACCGACACGCGGCAACGACCGCATCGAGGTGACGCGC 1380
Qy 1384 CTCAACGCGCACTTCCCGCAAGATCTCTGTTGCGGAGACCTTGGACGAGCCCCGAGCCATC 1443
Db 1381 CTCAACGCGCACTTCCCGCAAGATCTCTGTTGCGGAGACCTTGGACGAGCCCCGAGCCATC 1440
Qy 1444 GCATGTCACCCCGTGTGATGSGCTCATCTACTTGGACAGACTGGGAGAGAAACCTTAAATC 1503
Db 1441 GCATGTCACCCCGTGTGATGSGCTCATCTACTTGGACAGACTGGGAGAGAAACCTTAAATC 1500
Qy 1504 GAGTGTGCCAACTTTGGATGGGACGAGCGCGTGTGTGTTCAATGCTTCCCTCGGTTGG 1563
Db 1501 GAGTGTGCCAACTTTGGATGGGACGAGCGCGTGTGTGTTCAATGCTTCCCTCGGTTGG 1560
Qy 1564 CCCAACGCGCTTGGCCCTGGACCTTCAGAGGGGAGCTTACTTGGGAGAGCCCAAGACA 1623
Db 1561 CCCAACGCGCTTGGCCCTGGACCTTCAGAGGGGAGCTTACTTGGGAGAGCCCAAGACA 1620
Qy 1624 GACAAGATCGAGGTGATCAATGTTGATGGGACGAAGAGGGGGAACCTTCTTGGAGGAACAAG 1683
Db 1621 GACAAGATCGAGGTGATCAATGTTGATGGGACGAAGAGGGGGAACCTTCTTGGAGGAACAAG 1680
Qy 1684 CTCGCCGACATTTTTCGGGTTTCACTGCTGGGGGACTTCTATCTAATGAGACTGTGGCAG 1743
Db 1681 CTCGCCGACATTTTTCGGGTTTCACTGCTGGGGGACTTCTATCTAATGAGACTGTGGCAG 1740
Qy 1744 CGCGCAGCATCGAGCGGTTGCAAAAGTCAAGGCGCAGCGGGGAGCTCATCATTTGACCAG 1803
Db 1741 CGCGCAGCATCGAGCGGTTGCAAAAGTCAAGGCGCAGCGGGGAGCTCATCATTTGACCAG 1800
Qy 1804 CTGCGCGACCTGATGGGGCTCAAAAGCTGTGAATGTGGCCAAAGTCTGTGGAAACCAACCGC 1863
Db 1801 CTGCGCGACCTGATGGGGCTCAAAAGCTGTGAATGTGGCCAAAGTCTGTGGAAACCAACCGC 1860
Qy 1864 TGTGCGGACAGAAACGCGGGGTGAGCCACTGTGCTTCTTTCACACCCCAACGACCCG 1923
Db 1861 TGTGCGGACAGAAACGCGGGGTGAGCCACTGTGCTTCTTTCACACCCCAACGACCCG 1920
Qy 1924 TGTGCTGCCCATCGGCTGAGCTGTGAGTGCATGAAGACCTGTGATCTGTGCTGAG 1983
Db 1921 TGTGCTGCCCATCGGCTGAGCTGTGAGTGCATGAAGACCTGTGATCTGTGCTGAG 1980
Qy 1984 GCCTTCTTTGGTCTTTCACGACGAGCGCCCATCCACAGGATCTTCCCTCGAGACCAATTAAC 2043
Db 1981 GCCTTCTTTGGTCTTTCACGACGAGCGCCCATCCACAGGATCTTCCCTCGAGACCAATTAAC 2040
Qy 2044 AACGAGTGGGCGATCCCGCTCAGGGGCTCAAGGAGGCTCAGCCCTGGAATTTGATG 2103
Db 2041 AACGAGTGGGCGATCCCGCTCAGGGGCTCAAGGAGGCTCAGCCCTGGAATTTGATG 2100
Qy 2104 TCCAAACACCATCTACTTGGACAGAGCTCAGGCTGGAAGACCATCAGCGCGCTTCA 2163
Db 2101 TCCAAACACCATCTACTTGGACAGAGCTCAGGCTGGAAGACCATCAGCGCGCTTCA 2160
Qy 2164 AACGGGAGCTCGTGGAGCAGTGTGAGTTTGGCTTGAATACCCGAGGGCATGGCC 2223
Db 2161 AACGGGAGCTCGTGGAGCAGTGTGAGTTTGGCTTGAATACCCGAGGGCATGGCC 2220
Qy 2224 GTTGTGATGGGCAAGAACTTCTACTTGGGCGGACCTGGGACCAACAGAAATCCGAAGT 2283
Db 2221 GTTGTGATGGGCAAGAACTTCTACTTGGGCGGACCTGGGACCAACAGAAATCCGAAGT 2280
Qy 2284 GCGCGGCTGGACGGGCAAGTTCGCGCAAGTCTCTGTTGGAGGGACTTGGACAAACCGGAG 2343
Db 2281 GCGCGGCTGGACGGGCAAGTTCGCGCAAGTCTCTGTTGGAGGGACTTGGACAAACCGGAG 2340

QY 2344 TCGCTGGCCCTGGATCCCAACGAAGGGCTACATCTACTGACCGAGTGGGGGGCAAGCGG 2403
DB TCGCTGGCCCTGGATCCCAACGAAGGGCTACATCTACTGACCGAGTGGGGGGCAAGCGG 2400
QY 2404 AGGATCGTGGGGCTTTCATGACCGGACCAACTGCATGACGCTGGTGGACAAAGTGGGC 2463
DB AGGATCGTGGGGCTTTCATGACCGGACCAACTGCATGACGCTGGTGGACAAAGTGGGC 2460
QY 2464 CGGGCCAAAGCCTCACCATTTGACTACGCTGACACGAGCGCTTCTACTGACCGACCTGGAC 2523
DB CGGGCCAAAGCCTCACCATTTGACTACGCTGACACGAGCGCTTCTACTGACCGACCTGGAC 2520
QY 2524 ACCAATGATGAGTCTGCTGCAACATGCTGGGTGAGGAGCGGGTGGTGGTGGTGGTGGT 2583
DB ACCAATGATGAGTCTGCTGCAACATGCTGGGTGAGGAGCGGGTGGTGGTGGTGGTGGT 2580
QY 2584 CTCGCGCAACCGCTTTCGCTGACGAGTACGAGTACGAGTATATCTACTGACAGATGGAAT 2643
DB CTCGCGCAACCGCTTTCGCTGACGAGTACGAGTATATCTACTGACAGATGGAAT 2640
QY 2644 CTGACAGCATTTGAGCGGCGGACAAAGTACGAGCGGACCGGACCGCTCATCCAGGGC 2703
DB CTGACAGCATTTGAGCGGCGGACAAAGTACGAGCGGACCGGACCGCTCATCCAGGGC 2700
QY 2704 CACTGAGCTTCTGATGAGACATCTCTGCTGTTTCCACTCTCTCCGCGAGGATGGCTCAAT 2763
DB CACTGAGCTTCTGATGAGACATCTCTGCTGTTTCCACTCTCTCCGCGAGGATGGCTCAAT 2760
QY 2764 GACTGATGCAACAAACGGGCAAGTGTGGGCAAGTGTGCTTGGCATCCCGGGGGCCAC 2823
DB GACTGATGCAACAAACGGGCAAGTGTGGGCAAGTGTGCTTGGCATCCCGGGGGCCAC 2820
QY 2824 CGTGGGCTGGGCTTCTACATACACCTTGGACCCGAGCGGCGCACTGACGCGCGCC 2883
DB CGTGGGCTGGGCTTCTACATACACCTTGGACCCGAGCGGCGCACTGACGCGCGCC 2880
QY 2884 ACCACTTCTTGTCTGATGAGCAAAATCTGCGATCAGTCCGATGATCCGCGAGCACAG 2943
DB ACCACTTCTTGTCTGATGAGCAAAATCTGCGATCAGTCCGATGATCCGCGAGCACAG 2940
QY 2944 CACAGCCGGAATCTCATCTCTGCTTGCATGAGTGTGAGGAAAGTCAAAAGCCATGACTAT 3003
DB CACAGCCGGAATCTCATCTCTGCTTGCATGAGTGTGAGGAAAGTCAAAAGCCATGACTAT 3000
QY 3004 GACCCACTGGAAGATTTCTACTGAGTGTGAGGCGGCGGACCAATCAAGGAGCCAAAG 3063
DB GACCCACTGGAAGATTTCTACTGAGTGTGAGGCGGCGGACCAATCAAGGAGCCAAAG 3060
QY 3064 GACGACGGGACC-----CAGCCCTTTGTT 3087
DB GACGACGGGACC-----CAGCCCTTTGTT 3120
QY 3088 TTGACCTCTCTGAGCAAGGCCAAACCCAGACGAGCGCCCGCAGCTCAGCATCGAC 3147
DB TTGACCTCTCTGAGCAAGGCCAAACCCAGACGAGCGCCCGCAGCTCAGCATCGAC 3180
QY 3148 ATCTACGCGGACACTGTTCTGAGAGTGTGAGGCGGACCAATCAATCAAGTCCACAGG 3207
DB ATCTACGCGGACACTGTTCTGAGAGTGTGAGGCGGACCAATCAATCAAGTCCACAGG 3240
QY 3208 CTGAGCGGGGAAGCCATCGGGGTGTGCTGGGTGGGACCGGACCAAGCCAGGGCCATC 3267
DB CTGAGCGGGGAAGCCATCGGGGTGTGCTGGGTGGGACCGGACCAAGCCAGGGCCATC 3300
QY 3268 GTCTGTCAACCGGGAGCGAGGTACTCTGTACTTTACCAATGACGAGGACCGGGCGAGCCAAAG 3327
DB GTCTGTCAACCGGGAGCGAGGTACTCTGTACTTTACCAATGACGAGGACCGGGCGAGCCAAAG 3360
QY 3328 ATCGAACCGGACCGCTTGGACCGGACCGGAGGCTCTTTCACACCGCGGCTCATC 3387
DB ATCGAACCGGACCGCTTGGACCGGACCGGAGGCTCTTTCACACCGCGGCTCATC 3420
QY 3388 CGCCCTGTGGCCCTGGTGGTGGACAAACACATGAGGCAAGCTGTTCTGGGTGGACCGCGAC 3447

DB CGCCCTGTGGCCCTGGTGGTGGACAAACACATCTGGGCAAGCTGTTCTGGGTGGACCGCGAC 3421
QY 3448 CTGAAGCGCATTTGAGAGCTGTGACCTGTGACGGGGCAACCGCTGACCCCTGGAGACCGCC 3507
DB CTGAAGCGCATTTGAGAGCTGTGACCTGTGACGGGGCAACCGCTGACCCCTGGAGACCGCC 3540
QY 3508 AACATCTGTCAGCCTCTCTGGGCTGACCATCTCTGTCGAAGCATCTCTACTGGATCCACCGC 3567
DB AACATCTGTCAGCCTCTCTGGGCTGACCATCTCTGTCGAAGCATCTCTACTGGATCCACCGC 3600
QY 3568 CAGCAGCAGATGATCGAGCTGTGAGAGAGAACCCAGCGGGGCAACCGGACTCGCATCCAG 3627
DB CAGCAGCAGATGATCGAGCTGTGAGAGAGAACCCAGCGGGGCAACCGGACTCGCATCCAG 3660
QY 3628 GGCCTGTGGCCCACTCTCACTGGCATCCATGAGTGGAGGAGTCAAGCTGAGGAGGTTTC 3687
DB GGCCTGTGGCCCACTCTCACTGGCATCCATGAGTGGAGGAGTCAAGCTGAGGAGGTTTC 3720
QY 3688 TCAGCCCAACCATGTCGCCGTGACAAATGCTGCTGCTCCACATCTGTATTGCCAAGGTT 3747
DB TCAGCCCAACCATGTCGCCGTGACAAATGCTGCTGCTCCACATCTGTATTGCCAAGGTT 3780
QY 3748 GATGGGACACCAACCGGTGTCTATGCCCAGTCCACCTCTGCTCTCTGACAGAACCTGCTGACC 3807
DB GATGGGACACCAACCGGTGTCTATGCCCAGTCCACCTCTGCTCTCTGACAGAACCTGCTGACC 3840
QY 3808 TGTGGAGAGCGCCCACTCTGCTCCCGGACCAAGTGTGATGTCACAGGAGGAGATCGAC 3867
DB TGTGGAGAGCGCCCACTCTGCTCCCGGACCAAGTGTGATGTCACAGGAGGAGATCGAC 3900
QY 3868 TGTATCCCGGGGCTGGCGCTGTGACGGCTTTCCGAGTGGGATGACAGAGGACCGAG 3927
DB TGTATCCCGGGGCTGGCGCTGTGACGGCTTTCCGAGTGGGATGACAGAGGACCGAG 3960
QY 3928 GAGGCTTGCCTGCTGCTCCCGCGCCAGTTCCTCTCGCGCGGGGTGAGTGTGTGGAC 3987
DB GAGGCTTGCCTGCTGCTCCCGCGCCAGTTCCTCTCGCGCGGGGTGAGTGTGTGGAC 4020
QY 3988 CTGGCCTGTGCTGCGAGCGGACGAGACTGTGTGAGGACCGCTCAGACGAGGCGACTGT 4047
DB CTGGCCTGTGCTGCGAGCGGACGAGACTGTGTGAGGACCGCTCAGACGAGGCGACTGT 4080
QY 4048 GACGCCATCTGCTGCGCCCAACAGTCCGGTGTGGAGCGGCGGAGTGTCTCTATCAAA 4107
DB GACGCCATCTGCTGCGCCCAACAGTCCGGTGTGGAGCGGCGGAGTGTCTCTATCAAA 4140
QY 4108 CAGCAGTGCAGTCTCTTCCCGGACTGTATCGACGGCTCCGACGAGCTCATGTGTGAATC 4167
DB CAGCAGTGCAGTCTCTTCCCGGACTGTATCGACGGCTCCGACGAGCTCATGTGTGAATC 4200
QY 4168 ACCAAGCGCCCTCAGACGACAGCCCGGCGCCACAGAGTGCATCGGGCCCGCTCATTTGGC 4227
DB ACCAAGCGCCCTCAGACGACAGCCCGGCGCCACAGAGTGCATCGGGCCCGCTCATTTGGC 4260
QY 4228 ATCATCTCTCTCTTCTGCTCATGGGTGTGTTTATTTGTGTGTCAGCGGCTGTGTGTC 4287
DB ATCATCTCTCTCTTCTGCTCATGGGTGTGTTTATTTGTGTGTCAGCGGCTGTGTGTC 4320
QY 4288 CAGCGCTATGCGGGGGCCCAACCGGCCCTTCCCGCACAGTATGTTCAGCGGACCCCGCGAC 4347
DB CAGCGCTATGCGGGGGCCCAACCGGCCCTTCCCGCACAGTATGTTCAGCGGACCCCGCGAC 4380
QY 4348 GTGCCCTCTCAATTTTCATAGCCCGGGCGGTTTCCAGCATGCCCCCTTCCAGAGCATCGCA 4407
DB GTGCCCTCTCAATTTTCATAGCCCGGGCGGTTTCCAGCATGCCCCCTTCCAGAGCATCGCA 4440
QY 4408 TCGGGAAGTCCATGATGAGCTCCGTGAGCCTGATGCGGGGCGCGGGGCGGGGTGCCCTTG 4467
DB TCGGGAAGTCCATGATGAGCTCCGTGAGCCTGATGCGGGGCGCGGGGCGGGGTGCCCTTG 4500
QY 4468 TACGACCGGAAACACGCTCAGAGGGGCTCTGTCGACAGTCTGTCAGCACCAAGGCCACG 4527

Db 4501 TACGACCGGAACACGTCACAGGGGCTCTGTCAGCAGCTCGTCCAGCACGAAGGCCACG 4560
QY 4528 CTGTACCCGCGATCCTGAAACCGCGCGCTCCCGGCCACGGACCCCTCCCTGTACAAC 4587
Db 4561 CTGTACCCGCGATCCTGAAACCGCGCGCTCCCGGCCACGGACCCCTCCCTGTACAAC 4620
QY 4588 ATGGACATGTTCTACTCTTTCAAAACATTCGGGCCACTGCGAGACCGGTACAGGCCCTTACATC 4647
Db 4621 ATGGACATGTTCTACTCTTTCAAAACATTCGGGCCACTGCGAGACCGGTACAGGCCCTTACATC 4680
QY 4648 ATTCGAGGAATGGCGCCGCCCGACGACGCGCTTCAGCACCGCGTGTGTGACAGCGACTAC 4707
Db 4681 ATTCGAGGAATGGCGCCGCCCGACGACGCGCTTCAGCACCGCGTGTGTGACAGCGACTAC 4740
QY 4708 AGCGCCAGCCGCTGGAAGCCGAGCAAGTACTACCTGGATTGAACTCGGACTCAGACCCC 4767
Db 4741 AGCGCCAGCCGCTGGAAGCCGAGCAAGTACTACCTGGATTGAACTCGGACTCAGACCCC 4800
QY 4768 TATCCACCCCCACCCACGCCCCACAGCCAGTACTGTGCGGGAGGACAGCTGCCCGGCC 4827
Db 4801 TATCCACCCCCACCCACGCCCCACAGCCAGTACTGTGCGGGAGGACAGCTGCCCGGCC 4860
QY 4828 TCGCCCGCCACCGAGAGAGCTACTTCCATCTCTTCCCGCCCTCCGTCGCCCTGACG 4887
Db 4861 TCGCCCGCCACCGAGAGAGCTACTTCCATCTCTTCCCGCCCTCCGTCGCCCTGACG 4920
QY 4888 GACTCATCTGACCTCGCGCGGCCACCTCGGCTTCTGTGCCCCCTGTAATAGTTTTA 4947
Db 4921 GACTCATCTGACCTCGCGCGGCCACCTCGGCTTCTGTGCCCCCTGTAATAGTTTTA 4980
QY 4948 AATATGAACAAAGAAAAATATATTTATGATTTAAAAAATAAATAATTTGGGATTTT 5007
Db 4981 AATATGAACAAAGAAAAATATATTTATGATTTAAAAAATAAATAATTTGGGATTTT 5040
QY 5008 AAAAATGAGAAATGTGAACCTGTGATGGGTGGGCGAGGCTGGGAGAACTTTGTA 5063
Db 5041 AAAAATGAGAAATGTGAACCTGTGATGGGTGGGCGAGGCTGGGAGAACTTTGTA 5096

Search completed: February 18, 2005, 20:08:38
Job time : 21195 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 17, 2005, 01:25:15 ; Search time 49 seconds
(without alignments)
2460.372 Million cell updates/sec

Title: US-09-931-375A-2

Perfect score: 8740

Sequence: 1 MEAAPPPPPPLLLLLLLLL.....TERSYPFLPPPPSPCTDSS 1615

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA.*
- 1: /cgn2_6/ptodata/1/iaa/5A COMB.pap.*
 - 2: /cgn2_6/ptodata/1/iaa/5B COMB.pap.*
 - 3: /cgn2_6/ptodata/1/iaa/6A COMB.pap.*
 - 4: /cgn2_6/ptodata/1/iaa/6B COMB.pap.*
 - 5: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pap.*
 - 6: /cgn2_6/ptodata/1/iaa/backfiles1.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8736	100.0	1615	4	US-09-060-299-3
2	8736	100.0	1615	4	US-09-402-923A-3
3	8715	99.7	1615	4	US-09-544-398B-3
4	8715	99.7	1615	4	US-09-543-771B-3
5	8706	99.6	1615	4	US-09-544-398B-4
6	8706	99.6	1615	4	US-09-543-771B-4
7	8607	98.5	1591	4	US-09-060-299-4
8	8607	98.5	1591	4	US-09-402-923A-4
9	8607	98.5	1591	4	US-09-402-923A-43
10	8607	98.5	1591	4	US-09-060-299-29
11	8574	98.1	1639	4	US-09-402-923A-29
12	8574	98.1	1639	4	US-09-060-299-39
13	8568	98.0	1584	4	US-09-402-923A-39
14	8568	98.0	1584	4	US-09-060-299-42
15	8259.5	94.5	1614	4	US-09-060-299-42
16	8259.5	94.5	1614	4	US-09-402-923A-42
17	8208	93.9	1586	4	US-09-060-299-44
18	8208	93.9	1586	4	US-09-402-923A-44
19	7882	90.2	1451	4	US-09-060-299-25
20	7882	90.2	1451	4	US-09-402-923A-25
21	6153.5	70.4	1621	4	US-09-949-016-8450
22	5442	62.3	1148	4	US-09-949-016-8328
23	2937	33.6	550	4	US-09-060-299-7
24	2937	33.6	550	4	US-09-402-923A-7
25	2645.5	30.3	533	4	US-09-060-299-8
26	2645.5	30.3	533	4	US-09-402-923A-8
27	2193	25.1	4544	1	US-08-469-658-52

Sequence 52, Appl
Sequence 84, Appl
Sequence 84, Appl
Sequence 88, Appl
Sequence 90, Appl
Sequence 86, Appl
Sequence 8985, Ap
Sequence 3, Appl
Sequence 7, Appl
Sequence 40, Appl
Sequence 8313, Ap
Sequence 41833, A
Sequence 2, Appl
Sequence 2, Appl
Sequence 3, Appl
Sequence 3, Appl
Sequence 9528, Ap
Sequence 4, Appl

28 2193 25.1 4544 2 US-08-469-658-52
29 1876.5 21.5 4654 3 US-08-476-515A-84
30 1876.5 21.5 4655 3 US-08-652-877-84
31 1876.5 21.5 4655 3 US-08-652-877-88
32 1876.5 21.5 4655 3 US-08-652-877-90
33 1872.5 21.4 4655 3 US-08-652-877-86
34 1725.5 19.7 2362 4 US-09-949-016-8985
35 786.5 9.0 2213 1 US-08-727-034-3
36 774 8.9 2214 1 US-08-727-034-7
37 774 8.9 2214 4 US-09-919-039-40
38 694 7.9 1345 4 US-09-949-016-8313
39 688.5 7.9 396 4 US-09-270-767-41833
40 650 7.4 873 1 US-08-393-734-2
41 650 7.4 873 3 US-08-894-489-2
42 648 7.4 846 1 US-08-149-103-3
43 648 7.4 846 1 US-08-451-883-3
44 648 7.4 904 4 US-09-949-016-9528
45 646 7.4 846 1 US-08-149-103-4

ALIGNMENTS

RESULT 1
US-09-060-299-3
; Sequence 3, Application US/09060299
; Patent No. 6545137
; GENERAL INFORMATION:
; APPLICANT: Todd, John A
; APPLICANT: Hesse, John W
; APPLICANT: Caskey, Charles T
; APPLICANT: Cox, Roger D
; APPLICANT: Gerhold, David
; APPLICANT: Hammond, Holly
; APPLICANT: Hey, Patricia
; APPLICANT: Kawaguchi, Yoshihiko
; APPLICANT: Merriman, Tony R
; APPLICANT: Metzker, Michael L
; TITLE OF INVENTION: No. 6545137el Receptor
; NUMBER OF SEQUENCES: 455
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon and Vanderhye
; STREET: 1100 No. 6545137th Glebe Road, Eighth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: US
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/060,299
; FILING DATE: 15-APR-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/043,553
; FILING DATE: 15-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/048,740
; FILING DATE: 05-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: B.J.Sadoff
; REGISTRATION NUMBER: 36,663
; REFERENCE/DOCKET NUMBER: 620-35
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)816-4091
; TELEFAX: (703)816-4100
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1615 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

US-09-060-299-3

Query Match 100.0%; Score 8736; DB 4; Length 1615;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1614; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	1	MEAPPPPPWLLLLLLLLLALCCCPAPAAASPILLFANRRDRLVADGGVKLESTIVVS	60
Db	1	MEAPPPPPWLLLLLLLLLALCCCPAPAAASPILLFANRRDRLVADGGVKLESTIVVS	60
Qy	61	GLEDAAAVDFGSKGAVTWDVSEAIKQTYNOTGAOVNWTSGLVSPDGLACDWGK	120
Db	61	GLEDAAAVDFGSKGAVTWDVSEAIKQTYNOTGAOVNWTSGLVSPDGLACDWGK	120
Qy	121	KLYWTDSTNRIEVANLNGTSRKVLFWODLQDPAIALDPAHGYMYTWDGETPRIERAG	180
Db	121	KLYWTDSTNRIEVANLNGTSRKVLFWODLQDPAIALDPAHGYMYTWDGETPRIERAG	180
Qy	181	MDGSTRKLIIVDSIYWPNGLTIDLEEQKLYWADAKLSFIHRANLDGSPQKRVGSLTHP	240
Db	181	MDGSTRKLIIVDSIYWPNGLTIDLEEQKLYWADAKLSFIHRANLDGSPQKRVGSLTHP	240
Qy	241	FALTLSGDTLWTDQWTRSIHACNKRGTGKKEILSALYSPMDIQVLSQERQPPFHTRCE	300
Db	241	FALTLSGDTLWTDQWTRSIHACNKRGTGKKEILSALYSPMDIQVLSQERQPPFHTRCE	300
Qy	301	EDNGGCSHLCLLSPSEPPYTCACPTGVOQLQDNGRTCKAGAEVLLLARRTDLRISLDT	360
Db	301	EDNGGCSHLCLLSPSEPPYTCACPTGVOQLQDNGRTCKAGAEVLLLARRTDLRISLDT	360
Qy	361	DFTDVLQVDDIRHAIAIDYDPLSGYVYVTDDEVRATRRAYLDGSGAQTIVNTEINDPDG	420
Db	361	DFTDVLQVDDIRHAIAIDYDPLSGYVYVTDDEVRATRRAYLDGSGAQTIVNTEINDPDG	420
Qy	421	IADVARNLYTDTGTDRIEVTRNGTSRKILSVSEDLDEPRAIALHPVGLMTWTDGE	480
Db	421	IADVARNLYTDTGTDRIEVTRNGTSRKILSVSEDLDEPRAIALHPVGLMTWTDGE	480
Qy	481	NPKTECANLDQERRVLVNASLGNPGLALDQEGKLYWDAKTKIEVINVDGTGKRTL	540
Db	481	NPKTECANLDQERRVLVNASLGNPGLALDQEGKLYWDAKTKIEVINVDGTGKRTL	540
Qy	541	LEDKLPHFPGTLLGDFLYTWDQRRSIEVHKVKAASRDVLIIDOLPDLMLGLKAVNVAKV	600
Db	541	LEDKLPHFPGTLLGDFLYTWDQRRSIEVHKVKAASRDVLIIDOLPDLMLGLKAVNVAKV	600
Qy	601	GTNFCADRNGGCSHLCPFTPHATRCGCPIGLELLSDMKTCIVPEAFVFTSRAAIHRISL	660
Db	601	GTNFCADRNGGCSHLCPFTPHATRCGCPIGLELLSDMKTCIVPEAFVFTSRAAIHRISL	660
Qy	661	ETNNNDVAIPLTGVKEASALDFVSNHHIYWTDSVSLKTIISAPFNGSSVEHVVEFGLDYP	720
Db	661	ETNNNDVAIPLTGVKEASALDFVSNHHIYWTDSVSLKTIISAPFNGSSVEHVVEFGLDYP	720
Qy	721	EGMAVDWNGKLYWADGTGNRIEVARLDGQPROVLVWRDLNPRSLALDPTKGYIYWTEW	780
Db	721	EGMAVDWNGKLYWADGTGNRIEVARLDGQPROVLVWRDLNPRSLALDPTKGYIYWTEW	780
Qy	781	GGKPRIIVRAFMDGNTCMTLVKVGRANDLTIDYADQRLYWTDLTNMIESNNMLGQERVV	840
Db	781	GGKPRIIVRAFMDGNTCMTLVKVGRANDLTIDYADQRLYWTDLTNMIESNNMLGQERVV	840
Qy	841	IADDLPHFPGLTQYSDYIYWTDMNLHSIERADKTSGRNRTLIQGHDFVMDILVPHSSRQ	900
Db	841	IADDLPHFPGLTQYSDYIYWTDMNLHSIERADKTSGRNRTLIQGHDFVMDILVPHSSRQ	900
Qy	901	DGLNDCHNNQCCQOLCLAIIPGHRCCASHYTLDPSSRNCSPPTTLLFSQKSAISRMI	960
Db	901	DGLNDCHNNQCCQOLCLAIIPGHRCCASHYTLDPSSRNCSPPTTLLFSQKSAISRMI	960
Qy	961	PDDQHSPLILPLHGLNRVKAIDYDPLDKFTIYWDGQONIKRAKDDGTQPPVLTSLSQGG	1020
Db	961	PDDQHSPLILPLHGLNRVKAIDYDPLDKFTIYWDGQONIKRAKDDGTQPPVLTSLSQGG	1020

Qy	1021	NPDQPHDLSDIYSRTLFWTCEATNTINVHRLSGEAMGVVLRGDRDKPRAIVNAERGY	1080
Db	1021	NPDQPHDLSDIYSRTLFWTCEATNTINVHRLSGEAMGVVLRGDRDKPRAIVNAERGY	1080
Qy	1081	LYFTNMQDRAAKIERAALDGTREVLFTTGLIRVALVVDNTLGLFWVDADLKRIBSCD	1140
Db	1081	LYFTNMQDRAAKIERAALDGTREVLFTTGLIRVALVVDNTLGLFWVDADLKRIBSCD	1140
Qy	1141	LSGANRLTLEDANIVQPLGLTILGKLYWDRQOQMIERVEKTTGDKRTRIQGRVAHLTG	1200
Db	1141	LSGANRLTLEDANIVQPLGLTILGKLYWDRQOQMIERVEKTTGDKRTRIQGRVAHLTG	1200
Qy	1201	IHAVEVSLBEFSAHPCARDNGGCSHICIAKGDGTPRCSPVHLVLLQNLITCGEPPTCS	1260
Db	1201	IHAVEVSLBEFSAHPCARDNGGCSHICIAKGDGTPRCSPVHLVLLQNLITCGEPPTCS	1260
Qy	1261	PDQFACATGETDCIPGAWRCDGPFECDDQSEEGCPVCSAAQFFCARGQCCVDLRLRCDGE	1320
Db	1261	PDQFACATGETDCIPGAWRCDGPFECDDQSEEGCPVCSAAQFFCARGQCCVDLRLRCDGE	1320
Qy	1321	ADCCDRSDEADCAICLPNQFRCASGQCVLIKQCCDSFPDCIDGSDDELMECEITKPPSDDS	1380
Db	1321	ADCCDRSDEADCAICLPNQFRCASGQCVLIKQCCDSFPDCIDGSDDELMECEITKPPSDDS	1380
Qy	1381	PAHSSAIGPVIGIILSLFVNGGVYFVQVVCQRYAGANGPFPHEYVSGTBPVPLNFIAP	1440
Db	1381	PAHSSAIGPVIGIILSLFVNGGVYFVQVVCQRYAGANGPFPHEYVSGTBPVPLNFIAP	1440
Qy	1441	GGSHQGFPTGIACGKSNMSSVSLMGGRGVPLVDRNHVTGASSSSSSTKATLYPPLNP	1500
Db	1441	GGSHQGFPTGIACGKSNMSSVSLMGGRGVPLVDRNHVTGASSSSSSTKATLYPPLNP	1500
Qy	1501	PPSPATDPSLYNMDMYSSNIPATARYPIYIRGMAPPTTTCSTDVCDSDYSASRWKAS	1560
Db	1501	PPSPATDPSLYNMDMYSSNIPATARYPIYIRGMAPPTTTCSTDVCDSDYSASRWKAS	1560
Qy	1561	KYLDLNSDSDPYPPPTPHSQAEDSCPPSPATERSYHLPFPSPCTDSS	1615
Db	1561	KYLDLNSDSDPYPPPTPHSQAEDSCPPSPATERSYHLPFPSPCTDSS	1615

RESULT 2

US-09-402-923A-3
; Sequence 3, Application US/09402923A
; Patent No. 6555654

GENERAL INFORMATION:
APPLICANT: Todd, John A

Hess, John W
Caskey, Charles T
Cox, Roger D
Gerhold, David
Hammond, Holly
Hey, Patricia
Kawaguchi, Yoshihiko
Merriman, Tony R
Metzker, Michael L

TITLE OF INVENTION: No. 6555654el LDL-Receptor
NUMBER OF SEQUENCES: 455
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon and Vanderhye
STREET: 1100 No. 6555654th Glebe Road, Eighth Floor
CITY: Arlington
STATE: Virginia
COUNTRY: US
ZIP: VA 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/402,923A

FILING DATE: 14-Feb-2001
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB98/01102
FILING DATE: 15-APR-1998
APPLICATION NUMBER: US 60/043,553
FILING DATE: 15-APR-1997
APPLICATION NUMBER: US 60/048,740
FILING DATE: 05-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: B.J. Sadoff
REGISTRATION NUMBER: 36,663
REFERENCE/DOCKET NUMBER: 620-81
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4091
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1615 amino acids
TYPE: amino acid
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-402-923A-3

Query Match 100.0%; Score 8736; DB 4; Length 1615;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1614; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MEAAPPGPWPLLLLLLLLLLALCGCPAPAAASPLLLFANRRDVRLLVADAGGVKLESTIVVS 60
DB 1 MEAAPPGPWPLLLLLLLLLLALCGCPAPAAASPLLLFANRRDVRLLVADAGGVKLESTIVVS 60
QY 61 GLEDAAAVDFGSGAVTWDVSEBAIKQTYLNOTGAQVNVVSLVSPGLACDWGK 120
DB 61 GLEDAAAVDFGSGAVTWDVSEBAIKQTYLNOTGAQVNVVSLVSPGLACDWGK 120
QY 121 KLYWTDSETNRIEVANLNGTSRKVLFWQDLQPAIALDPAGHYMYWTDWGETPRIERAG 180
DB 121 KLYWTDSETNRIEVANLNGTSRKVLFWQDLQPAIALDPAGHYMYWTDWGETPRIERAG 180
QY 181 MDGSTRKLIIVDSIYWPNGLIIDLEQKLYWADAKLSFIHRANLDGSPRQKVVGSLTLP 240
DB 181 MDGSTRKLIIVDSIYWPNGLIIDLEQKLYWADAKLSFIHRANLDGSPRQKVVGSLTLP 240
QY 241 FALTLSGDTLWYTDWQTRSIIHACNKRGTGKKEILSALYSFMDIQVLQSOERQPFHTNCE 300
DB 241 FALTLSGDTLWYTDWQTRSIIHACNKRGTGKKEILSALYSFMDIQVLQSOERQPFHTNCE 300
QY 301 EDNGGCSHLCLLSPSEPPYTCACPTGVQLQDNGRTCKAGAEVLLARRTDLRISLDT 360
DB 301 EDNGGCSHLCLLSPSEPPYTCACPTGVQLQDNGRTCKAGAEVLLARRTDLRISLDT 360
QY 361 DFTDIVLQDDIRHAIAIDYDPLEGYVYVWTDDEVRAIRRAYLDGSGAQLVNTINPDG 420
DB 361 DFTDIVLQDDIRHAIAIDYDPLEGYVYVWTDDEVRAIRRAYLDGSGAQLVNTINPDG 420
QY 421 IAVDWARNLWYTDGTDRIEVTRNGTSRKILVSEDLDEPRAIALHPVGMATWNGE 480
DB 421 IAVDWARNLWYTDGTDRIEVTRNGTSRKILVSEDLDEPRAIALHPVGMATWNGE 480
QY 481 NPKIECANLDQERRVLYNLSGPNGLALDQEGKLYWDAKTDKIEVINVDGTKRRTL 540
DB 481 NPKIECANLDQERRVLYNLSGPNGLALDQEGKLYWDAKTDKIEVINVDGTKRRTL 540
QY 541 LEDKLPHFGLTLDGDFIYVWTDWQRRSIEVHVKKASRDVLIIDQLPDLMLGLKAVNAKV 600
DB 541 LEDKLPHFGLTLDGDFIYVWTDWQRRSIEVHVKKASRDVLIIDQLPDLMLGLKAVNAKV 600
QY 601 GTNPCADNGGCSHLCPFTPHATCGCPIGLELSDMKTCIVPRAFLVFTSRAIHRISL 660
DB 601 GTNPCADNGGCSHLCPFTPHATCGCPIGLELSDMKTCIVPRAFLVFTSRAIHRISL 660
QY 661 ETNNNDVAIPTGVKEASALDFVSNHHIYWTDVSLKTIISRAFMNGSSVEHVFGLDYP 720
DB 661 ETNNNDVAIPTGVKEASALDFVSNHHIYWTDVSLKTIISRAFMNGSSVEHVFGLDYP 720

DB 661 ETNNNDVAIPTGVKEASALDFVSNHHIYWTDVSLKTIISRAFMNGSSVEHVFGLDYP 720
QY 721 EGMAVDWMGKULYWADTGTNRRIEVARLDGQPROVLVWRDLNPRSLALDPTKGYIYWTEW 780
DB 721 EGMAVDWMGKULYWADTGTNRRIEVARLDGQPROVLVWRDLNPRSLALDPTKGYIYWTEW 780
QY 781 GKGPRIVRAFMWDGTCMTLVDKVGRANDLTIDYADQRLYWTDLTNMIENSNMLGQERVV 840
DB 781 GKGPRIVRAFMWDGTCMTLVDKVGRANDLTIDYADQRLYWTDLTNMIENSNMLGQERVV 840
QY 841 IADLPHPPGLTQYSYVYVWTDWNLHLSIERADKTSGRNRTLIQGHLDVFMVDILVFHSSRQ 900
DB 841 IADLPHPPGLTQYSYVYVWTDWNLHLSIERADKTSGRNRTLIQGHLDVFMVDILVFHSSRQ 900
QY 901 DGLNDCHNNGCCGQLCLAIIPGHRCCGASHYTLDPSSRNCSPPTTFLFKQSAISRM 960
DB 901 DGLNDCHNNGCCGQLCLAIIPGHRCCGASHYTLDPSSRNCSPPTTFLFKQSAISRM 960
QY 961 PDDQHSPLILPLHLNRNKAIDYDPLDKFYVWDGRONIKRAKDDGTQPPVLTSLSQO 1020
DB 961 PDDQHSPLILPLHLNRNKAIDYDPLDKFYVWDGRONIKRAKDDGTQPPVLTSLSQO 1020
QY 1021 NPDRQPHDLSIDIYSRTLFWTCEATNTINVRHLSGEAMGVVLRGDRDKPRAIVVNAERG 1080
DB 1021 NPDRQPHDLSIDIYSRTLFWTCEATNTINVRHLSGEAMGVVLRGDRDKPRAIVVNAERG 1080
QY 1081 LYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLFWVDADLKRIESCD 1140
DB 1081 LYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLFWVDADLKRIESCD 1140
QY 1141 LSGANRLTLEDANIVQPLGLTILGKHLWIDRQOQMIEVERKTTGDKRTRIQGRVAHLTG 1200
DB 1141 LSGANRLTLEDANIVQPLGLTILGKHLWIDRQOQMIEVERKTTGDKRTRIQGRVAHLTG 1200
QY 1201 IHAVEEVSLEEFSAHPCARDNGGCSHICIAKGDGTPRCSPVHLVLLQNLLTCGEPPTCS 1260
DB 1201 IHAVEEVSLEEFSAHPCARDNGGCSHICIAKGDGTPRCSPVHLVLLQNLLTCGEPPTCS 1260
QY 1261 PDQACATGETDCIPGAWRCDDGPECDSDDEECPCVCSAAQFCARGCQVDBLRCDGE 1320
DB 1261 PDQACATGETDCIPGAWRCDDGPECDSDDEECPCVCSAAQFCARGCQVDBLRCDGE 1320
QY 1321 ADCQDRDEADCDAILCPNQFRCASGCVLTKQCDSPDPCIDGSDLMCEITKPPSDDS 1380
DB 1321 ADCQDRDEADCDAILCPNQFRCASGCVLTKQCDSPDPCIDGSDLMCEITKPPSDDS 1380
QY 1381 PAHSSAIGPVIGIILSLFVMGVYFVQVVRVQRYAGANGPPHYVSGTTPHPLNFIAP 1440
DB 1381 PAHSSAIGPVIGIILSLFVMGVYFVQVVRVQRYAGANGPPHYVSGTTPHPLNFIAP 1440
QY 1441 GGSQHGPTGIACGKSNMSSVSLMGGRGVPLVDRNHTVGTASSSSSTKATLPPILNP 1500
DB 1441 GGSQHGPTGIACGKSNMSSVSLMGGRGVPLVDRNHTVGTASSSSSTKATLPPILNP 1500
QY 1501 PPSPATPSLYNMDFYSSNIPATVPYPIIRGMAPPPTPCSTDVCDSDSYASRKAAS 1560
DB 1501 PPSPATPSLYNMDFYSSNIPATVPYPIIRGMAPPPTPCSTDVCDSDSYASRKAAS 1560
QY 1561 KYLDLNSDSDPYPPPTPHSQALESADSCPPSPATERSYFHLFPSPCTDSS 1615
DB 1561 KYLDLNSDSDPYPPPTPHSQALESADSCPPSPATERSYFHLFPSPCTDSS 1615

RESULT 3
US-09-544-398B-3
; Sequence 3, Application US/09544398B
; Patent No. 6770461
; GENERAL INFORMATION:
; APPLICANT: Carulli, John P.
; APPLICANT: Little, Randall D.
; APPLICANT: Recker, Robert R.
; APPLICANT: Johnson, Mark L.
; TITLE OF INVENTION: High bone mass gene of 11q13.3

```
; FILE REFERENCE: 032796-013
; CURRENT APPLICATION NUMBER: US/09/544,398B
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: US/09/229,319
; PRIOR FILING DATE: 1999-01-13
; PRIOR APPLICATION NUMBER: US/09/071,449
; PRIOR FILING DATE: 1998-01-13
; PRIOR APPLICATION NUMBER: US/09/105,511
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 641
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1615
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-544-398B-3

Query Match          99.7%; Score 8715; DB 4; Length 1615;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1610; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MEAAPGPPWPLLLLLLLLLLALCGCPAPAAAPLFFANRRDRLVADAGGVKLESTIVVS 60
DB 1 MEAAPGPPWPLLLLLLLLLLALCGCPAPAAAPLFFANRRDRLVADAGGVKLESTIVVS 60
QY 61 GLEDAAAVDFQSKGAVYTDVSEEAIKQTYLNOTGAQVQNVVIGSLVSPDGLACDWGK 120
DB 61 GLEDAAAVDFQSKGAVYTDVSEEAIKQTYLNOTGAQVQNVVIGSLVSPDGLACDWGK 120
QY 121 KLYWTDSETNREIVANLNGTSKVLFWQDLOPRAIALDPAGHYMYMTDGETPRIERAG 180
DB 121 KLYWTDSETNREIVANLNGTSKVLFWQDLOPRAIALDPAGHYMYMTDGETPRIERAG 180
QY 181 MDGSTRKIIVSDIYWPNGLTIDLEQKLYWADAKLSFIHRANLDGSPQKVVESGLTHP 240
DB 181 MDGSTRKIIVSDIYWPNGLTIDLEQKLYWADAKLSFIHRANLDGSPQKVVESGLTHP 240
QY 241 FALTLSGDTLYWTDWQTSIHACNKRKTKGKKEILSALYSMPDIQVLQSOERQPPHTTCE 300
DB 241 FALTLSGDTLYWTDWQTSIHACNKRKTKGKKEILSALYSMPDIQVLQSOERQPPHTTCE 300
QY 301 EDNGGCSHLLSPSEPFYTCACPTGVQDNGRTCKAGAEVLLARLRLSLDTP 360
DB 301 EDNGGCSHLLSPSEPFYTCACPTGVQDNGRTCKAGAEVLLARLRLSLDTP 360
QY 361 DFTDIVLQVDDIRHAIAIDYDPLEGVYVWTDDEVRAIRRAYLDGSGAQTLYNTEINDPDG 420
DB 361 DFTDIVLQVDDIRHAIAIDYDPLEGVYVWTDDEVRAIRRAYLDGSGAQTLYNTEINDPDG 420
QY 421 IAVDVARNLYWTDGTDRIEVTRNGTSKTLVSEDLDEPRALALHPVMGLMYWTDWGE 480
DB 421 IAVDVARNLYWTDGTDRIEVTRNGTSKTLVSEDLDEPRALALHPVMGLMYWTDWGE 480
QY 481 NPKIECANLDGQERRVIVNNSALGWPNGLALDIQEGKLYWDAKTDKIEVINVDGTKRRTL 540
DB 481 NPKIECANLDGQERRVIVNNSALGWPNGLALDIQEGKLYWDAKTDKIEVINVDGTKRRTL 540
QY 541 LEDKLPHPFGLTLDGFIYWTWQRRSIRVHKVKAARDVVIDQLPDLMLKAVNVAKV 600
DB 541 LEDKLPHPFGLTLDGFIYWTWQRRSIRVHKVKAARDVVIDQLPDLMLKAVNVAKV 600
QY 601 GTNPCCADRGCGSHLCFFTPHATRCGCPGLELLSDMTKCIVPEAFVFTSRAAIAHRLSL 660
DB 601 GTNPCCADRGCGSHLCFFTPHATRCGCPGLELLSDMTKCIVPEAFVFTSRAAIAHRLSL 660
QY 661 ETNNNDVAIPLTGKVEASALDFVSNHHIYWTDVSLKTSIRAFMNGSSVEHVVEFLDYP 720
DB 661 ETNNNDVAIPLTGKVEASALDFVSNHHIYWTDVSLKTSIRAFMNGSSVEHVVEFLDYP 720
QY 721 EGMADVWGMKNLYWADTGTNRLEVARLDGQFQVLRDLNPRSLALDPTKGYIYWTEW 780
DB 721 EGMADVWGMKNLYWADTGTNRLEVARLDGQFQVLRDLNPRSLALDPTKGYIYWTEW 780
```

```
QY 781 GGPRIIVRAFMGDTNCMTLVLDKVGANDLTIDVADORLYWTDLTDTNNIESNNMLGQSRVV 840
DB 781 GGPRIIVRAFMGDTNCMTLVLDKVGANDLTIDVADORLYWTDLTDTNNIESNNMLGQSRVV 840
QY 841 IADDLPHFPGLTQYSDIYIYWTDMNLHSIERADKTSGRNRTLIOGHLDFVMDILVFHSSRQ 900
DB 841 IADDLPHFPGLTQYSDIYIYWTDMNLHSIERADKTSGRNRTLIOGHLDFVMDILVFHSSRQ 900
QY 901 DGLNDCMHNNGCQGLCIAIPGGHRCGCASHYTLDPSSRNCSPPTTLLFSQKSAISRM 960
DB 901 DGLNDCMHNNGCQGLCIAIPGGHRCGCASHYTLDPSSRNCSPPTTLLFSQKSAISRM 960
QY 961 PDDQHSPLILPLHGLRNVAIDYDPLDKFYWVDRQNIKRAKDDGTQFVLTSLSQGQ 1020
DB 961 PDDQHSPLILPLHGLRNVAIDYDPLDKFYWVDRQNIKRAKDDGTQFVLTSLSQGQ 1020
QY 1021 NPDQPHDLSIDIYSRTLFWTCEATNTINVHRLSGEAMGVVLRGDRDKPRAIVVNAERG 1080
DB 1021 NPDQPHDLSIDIYSRTLFWTCEATNTINVHRLSGEAMGVVLRGDRDKPRAIVVNAERG 1080
QY 1081 LYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLKFWVDADLAKRIESC 1140
DB 1081 LYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLKFWVDADLAKRIESC 1140
QY 1141 LSGANRLTLEDAIVQPLGLTILGKHYWIDRQOQMIERVEKTTGDKRTRIQRVAHLTG 1200
DB 1141 LSGANRLTLEDAIVQPLGLTILGKHYWIDRQOQMIERVEKTTGDKRTRIQRVAHLTG 1200
QY 1201 IHAVERVSLSEFSAHPCARDNGGCSHICIAKGDGTGPRCSPVHLVLLQNLITCGEPPTCS 1260
DB 1201 IHAVERVSLSEFSAHPCARDNGGCSHICIAKGDGTGPRCSPVHLVLLQNLITCGEPPTCS 1260
QY 1261 PDQFACATGEDIICPGAWRCDDQDEECPCVCSAAQFPCARGQCVDLRLRCDGE 1320
DB 1261 PDQFACATGEDIICPGAWRCDDQDEECPCVCSAAQFPCARGQCVDLRLRCDGE 1320
QY 1321 ADCQDRDEADCDACILPNQFRFCASGQCVLIKQCCDSFPDCIDGSDBLMCEITKPPSDDS 1380
DB 1321 ADCQDRDEADCDACILPNQFRFCASGQCVLIKQCCDSFPDCIDGSDBLMCEITKPPSDDS 1380
QY 1381 PAHSSAIGPVIGIILSLFVMGGVYFVQVVCQRYAGANGPPHYYSGTTPHPLNFAP 1440
DB 1381 PAHSSAIGPVIGIILSLFVMGGVYFVQVVCQRYAGANGPPHYYSGTTPHPLNFAP 1440
QY 1441 GSGHQGPFTGIACGKSMSSVSLMGGRGVPLVYDRNHVTGASSSSSSSTKATLYPPLNP 1500
DB 1441 GSGHQGPFTGIACGKSMSSVSLMGGRGVPLVYDRNHVTGASSSSSSSTKATLYPPLNP 1500
QY 1501 PPSPATDPSLYNMDMYSSNIPATARPYPYIIIRGMAPPTTPTCSTDVCDSDYSASRWKAS 1560
DB 1501 PPSPATDPSLYNMDMYSSNIPATARPYPYIIIRGMAPPTTPTCSTDVCDSDYSASRWKAS 1560
QY 1561 KYIYDLNDSDDPYPPPTPHSHOYLSEAEDSCPPSPATERSYFHLFPFPPSPCTDSS 1615
DB 1561 KYIYDLNDSDDPYPPPTPHSHOYLSEAEDSCPPSPATERSYFHLFPFPPSPCTDSS 1615
```

RESULT 4

```
US-09-543-771B-3
; Sequence 3, Application US/09543771B
; Patent No. 6780609
; GENERAL INFORMATION:
; APPLICANT: Carulli, John P.
; APPLICANT: Little, Randall D.
; APPLICANT: Recker, Robert L.
; APPLICANT: Johnson, Mark L.
; TITLE OF INVENTION: High bone mass gene of 11q13.3
; FILE REFERENCE: 032796-014
; CURRENT APPLICATION NUMBER: US/09/543,771B
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 09/229,319
; PRIOR FILING DATE: 1999-01-13
; PRIOR APPLICATION NUMBER: US 60/071,449
```

;
; PRIOR FILING DATE: 1998-01-13
; PRIOR APPLICATION NUMBER: US 60/105,511
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 641
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1615
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-543-771B-3

Query Match 99.7%; Score 8715; DB 4; Length 1615;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1610; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MEAAPPGPWLLLLLLLLLALCGCPAPAAASPLLLFANRRDVRVLDAGGVKLESTIVS 60
DB 1 MEAAPPGPWLLLLLLLLLALCGCPAPAAASPLLLFANRRDVRVLDAGGVKLESTIVS 60

QY 61 GLEDAAAVDFQFSGAVVYTDVSEAIKQTYLNOTGAQVNVVLSGLVSPDGLACDWGK 120
DB 61 GLEDAAAVDFQFSGAVVYTDVSEAIKQTYLNOTGAQVNVVLSGLVSPDGLACDWGK 120

QY 121 KLYWTDSTNRIEVANLNGTSRKVLFWQDLQDPAIALDPAHGYMYWTDWGETPRIERAG 180
DB 121 KLYWTDSTNRIEVANLNGTSRKVLFWQDLQDPAIALDPAHGYMYWTDWGETPRIERAG 180

QY 181 MDGSTRKLIIVSDIYNPLGLTIDLEOKLYWADAKLSFIHRANLDGSPKQKVGSLTLP 240
DB 181 MDGSTRKLIIVSDIYNPLGLTIDLEOKLYWADAKLSFIHRANLDGSPKQKVGSLTLP 240

QY 241 FALTLSGDTLWYDQWTSIHACNKRKTCGKKEILSALYSMDIQVLSQERQPPFHTRCE 300
DB 241 FALTLSGDTLWYDQWTSIHACNKRKTCGKKEILSALYSMDIQVLSQERQPPFHTRCE 300

QY 301 EDNGGCSHLCLLSPSEPYTCACPTGVQVQDNGRTCKAGAEVLLARRTDLRRISLDP 360
DB 301 EDNGGCSHLCLLSPSEPYTCACPTGVQVQDNGRTCKAGAEVLLARRTDLRRISLDP 360

QY 361 DFTDIVLQVDDIRHAIAIDYDPLEGVYVYTDDEVRALREAYLDGSGAQLVNTENDPDG 420
DB 361 DFTDIVLQVDDIRHAIAIDYDPLEGVYVYTDDEVRALREAYLDGSGAQLVNTENDPDG 420

QY 421 IAVDWARNLYWTDGTDRIEVRLNGTSRKILVSEDLDEPRALHPVGMGLMYWTDWGE 480
DB 421 IAVDWARNLYWTDGTDRIEVRLNGTSRKILVSEDLDEPRALHPVGMGLMYWTDWGE 480

QY 481 NPKIECANLDGOERRVLYNASIGWPNGLALDLEQKLYWDAKTDKIEVINVDGTKRRTL 540
DB 481 NPKIECANLDGOERRVLYNASIGWPNGLALDLEQKLYWDAKTDKIEVINVDGTKRRTL 540

QY 541 LEDKLPHIFGFTLIGDFFIYTDWQRRSIRVHKVKASRDVLIIDQLPDLMLKANVAKV 600
DB 541 LEDKLPHIFGFTLIGDFFIYTDWQRRSIRVHKVKASRDVLIIDQLPDLMLKANVAKV 600

QY 601 GTNPCADRNGGCSHLCPFTPHATRCGCPGLELLSDMKTICVPEAFVFTSRAAHRISL 660
DB 601 GTNPCADRNGGCSHLCPFTPHATRCGCPGLELLSDMKTICVPEAFVFTSRAAHRISL 660

QY 661 ETNNNDVAIPITGVKEASALDFDVSNHHIYWDVSLKTIISRAFWNGSSVEHVFGLDYP 720
DB 661 ETNNNDVAIPITGVKEASALDFDVSNHHIYWDVSLKTIISRAFWNGSSVEHVFGLDYP 720

QY 721 EGMADVWNGKLYWADTGTNRIEVARLDGQFQVLYWRDLNPRSLALDPTKGYIYTEW 780
DB 721 EGMADVWNGKLYWADTGTNRIEVARLDGQFQVLYWRDLNPRSLALDPTKGYIYTEW 780

QY 781 GKGKRIVRAFMDGTNCMTLVKVGRLDITDADQRLYTDLDTNMIESNNMLGOERVV 840
DB 781 GKGKRIVRAFMDGTNCMTLVKVGRLDITDADQRLYTDLDTNMIESNNMLGOERVV 840

QY 841 IADDLPHFPGLTQYSDYIYTDWNLHSIERADKTSGRNRTLIQGHLDPFVMDILVFHSSRQ 900
DB 841 IADDLPHFPGLTQYSDYIYTDWNLHSIERADKTSGRNRTLIQGHLDPFVMDILVFHSSRQ 900

RESULT 5
US-09-544-398B-4
; Sequence 4, Application US/09544398B
; Patent No. 6770461
; GENERAL INFORMATION:
; APPLICANT: Carulli, John P.
; APPLICANT: Little, Randall D.
; APPLICANT: Recker, Robert R.
; APPLICANT: Johnson, Mark L.
; TITLE OF INVENTION: High bone mass gene of 11q13.3
; FILE REFERENCE: 032796-013
; CURRENT APPLICATION NUMBER: US/09/544.398B
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: US 09/229,319
; PRIOR FILING DATE: 1999-01-13
; PRIOR APPLICATION NUMBER: US 60/071,449
; PRIOR FILING DATE: 1998-01-13
; PRIOR APPLICATION NUMBER: US 60/105,511
; NUMBER OF SEQ ID NOS: 641
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4

```
; LENGTH: 1615
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-544-398B-4

Query Match      99.6%; Score 8706; DB 4; Length 1615;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1609; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 MEAAPGPPWPLLLLLLLLLLALCGCPAPAAASPLLLFANRRDVRIVLDAGGVKLESTIVVS 60
DB 1 MEAAPGPPWPLLLLLLLLLLALCGCPAPAAASPLLLFANRRDVRIVLDAGGVKLESTIVVS 60

QY 61 GLEDAAAVDFQSKGAVYWDVSEBAIKQTYLNOTGAAYQNVVVISGLVSPDGLACDWYVK 120
DB 61 GLEDAAAVDFQSKGAVYWDVSEBAIKQTYLNOTGAAYQNVVVISGLVSPDGLACDWYVK 120

QY 121 KLYWTDSETNRIEVANLNGTSKVLFWQDLDOPRAIALDPAGHYMYWTDWGETPRIERAG 180
DB 121 KLYWTDSETNRIEVANLNGTSKVLFWQDLDOPRAIALDPAGHYMYWTDWGETPRIERAG 180

QY 181 MDGSTRKIIVSDIYWPNGLTIDLEQKLYWADAKLSPIHRANLDGSPQKVVESGLTTP 240
DB 181 MDGSTRKIIVSDIYWPNGLTIDLEQKLYWADAKLSPIHRANLDGSPQKVVESGLTTP 240

QY 241 FALTLSGDTLYWTDQWOTSIHACNKRITGGRKEILSALYSPMDIQVLQOERQPPFHTTCE 300
DB 241 FALTLSGDTLYWTDQWOTSIHACNKRITGGRKEILSALYSPMDIQVLQOERQPPFHTTCE 300

QY 301 EDNGGSHLCLLSPSEPYTCACPTGVQVQNGRTCKAGABEVLLIARRTDLRISLDT 360
DB 301 EDNGGSHLCLLSPSEPYTCACPTGVQVQNGRTCKAGABEVLLIARRTDLRISLDT 360

QY 361 DFTDVLQVDDIRHAIADYDPLEGYVYWDDEVRAIRRAYLDGSGAOTLVNTEINDPDG 420
DB 361 DFTDVLQVDDIRHAIADYDPLEGYVYWDDEVRAIRRAYLDGSGAOTLVNTEINDPDG 420

QY 421 IAVDWARNLYWTDGTDRIEVTRNGTSRKILYSEDLDEPRATALHPVGMGLMYWTDWGE 480
DB 421 IAVDWARNLYWTDGTDRIEVTRNGTSRKILYSEDLDEPRATALHPVGMGLMYWTDWGE 480

QY 481 NPKIECANLDQERRVLVNASLWPNGLALDQEGKLYWDAKTDKIEVINVDGTKRRTL 540
DB 481 NPKIECANLDQERRVLVNASLWPNGLALDQEGKLYWDAKTDKIEVINVDGTKRRTL 540

QY 541 LEDKLPHIFGTLGDFIYTDWQRRSIEVHKVKASRDVIIDQLPDLMLGLKAVNVAKV 600
DB 541 LEDKLPHIFGTLGDFIYTDWQRRSIEVHKVKASRDVIIDQLPDLMLGLKAVNVAKV 600

QY 601 GTNPCADRNCGCSHLCTFTPHATRCGCPIGLELLSDMKTCIVPEAFIVFTSRAAHRISL 660
DB 601 GTNPCADRNCGCSHLCTFTPHATRCGCPIGLELLSDMKTCIVPEAFIVFTSRAAHRISL 660

QY 661 ETNNNDVAIPLTGVKEASALDFVSNHHIYWDVSLKTIISRAFMNGSSVEHVVFGLDYP 720
DB 661 ETNNNDVAIPLTGVKEASALDFVSNHHIYWDVSLKTIISRAFMNGSSVEHVVFGLDYP 720

QY 721 EGMVDMWNGKLYADTGTNRIEVARLDQQRQVLRWDLNPRSLALDPTKGYIYTWTEW 780
DB 721 EGMVDMWNGKLYADTGTNRIEVARLDQQRQVLRWDLNPRSLALDPTKGYIYTWTEW 780

QY 781 GGGPRIVRAFMDGTNCMTLVKVGANDLTIDYADQRLYTDLDTNMISSNMLQGERVV 840
DB 781 GGGPRIVRAFMDGTNCMTLVKVGANDLTIDYADQRLYTDLDTNMISSNMLQGERVV 840

QY 841 IADDLPHFPGLTQYSDIYWTWNLHLSIERADKTSGRNRTLIQGHDFWMDILVFHSSRQ 900
DB 841 IADDLPHFPGLTQYSDIYWTWNLHLSIERADKTSGRNRTLIQGHDFWMDILVFHSSRQ 900

QY 901 DGLNDGMHNNCGGOLCLAIPEGHRCGCASHYTLDPSSRNCSPPTTLLFSQKSAISRM 960
DB 901 DGLNDGMHNNCGGOLCLAIPEGHRCGCASHYTLDPSSRNCSPPTTLLFSQKSAISRM 960

QY 961 PDQHSDDLILPLHCLRNKKAIDYDPLDKFYWVDGRQNIKRAKXDDGTQPPVLTSLSQG 1020
DB 961 PDQHSDDLILPLHCLRNKKAIDYDPLDKFYWVDGRQNIKRAKXDDGTQPPVLTSLSQG 1020

QY 1021 NPDQPHDLSIDIYSRTLFWTCEATNTINVRHLSGEAMGVVLRGDRDKPRAIVVNAERG 1080
DB 1021 NPDQPHDLSIDIYSRTLFWTCEATNTINVRHLSGEAMGVVLRGDRDKPRAIVVNAERG 1080

QY 1081 LYFTNQMDRAAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLFWVDADLKRIESC 1140
DB 1081 LYFTNQMDRAAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLFWVDADLKRIESC 1140

QY 1141 LSGANRLTLEDAIVOPGLTILGKHLIYWDROOQMIERVEKTTGDKRTRIQGRVAHLTG 1200
DB 1141 LSGANRLTLEDAIVOPGLTILGKHLIYWDROOQMIERVEKTTGDKRTRIQGRVAHLTG 1200

QY 1201 IHAVEVSLBEFSAHPCARNDGGCSHICIAKGDGTPRCSCPVLHVLVLLQNLITCGEPPTCS 1260
DB 1201 IHAVEVSLBEFSAHPCARNDGGCSHICIAKGDGTPRCSCPVLHVLVLLQNLITCGEPPTCS 1260

QY 1261 PDQFACATGEIDCIPGAWRCDGPFECDDQSDDEGCPVCSAAQFPCARGQCVDLRLRCDGE 1320
DB 1261 PDQFACATGEIDCIPGAWRCDGPFECDDQSDDEGCPVCSAAQFPCARGQCVDLRLRCDGE 1320

QY 1321 ADCQDRSDEADCDALCLPNQFRCSGOCVLIKQOCDFPDICIDGSDLMCEITKPPSDDS 1380
DB 1321 ADCQDRSDEADCDALCLPNQFRCSGOCVLIKQOCDFPDICIDGSDLMCEITKPPSDDS 1380

QY 1381 PAHSSAIGPVGIIILSLFVWGGVYFVQVRVVCQRYAGANGPPHEVYSGTTPHVPINFIAP 1440
DB 1381 PAHSSAIGPVGIIILSLFVWGGVYFVQVRVVCQRYAGANGPPHEVYSGTTPHVPINFIAP 1440

QY 1441 GGSQHGFTGTIACGKSMSSVSLMGGRGVPLYDRNHVHTGASSSSSSSTKATLPPILNP 1500
DB 1441 GGSQHGFTGTIACGKSMSSVSLMGGRGVPLYDRNHVHTGASSSSSSSTKATLPPILNP 1500

QY 1501 PPSPATPSLYNMDFYSSNIPATAPRVPYIIRGMAPPPTPCSTDVDCSDYSASRWKAS 1560
DB 1501 PPSPATPSLYNMDFYSSNIPATAPRVPYIIRGMAPPPTPCSTDVDCSDYSASRWKAS 1560

QY 1561 KYLDLNSDSDPYPPPTPHSPHYSQYLSAEDSCPPSPATERSYFHLFPFPPSPCTDSS 1615
DB 1561 KYLDLNSDSDPYPPPTPHSPHYSQYLSAEDSCPPSPATERSYFHLFPFPPSPCTDSS 1615

RESULT 6
US-09-543-771B-4
; Sequence 4, Application US/09543771B
; Patent No. 6780609
; GENERAL INFORMATION:
; APPLICANT: Carulli, John P.
; APPLICANT: Little, Randall D.
; APPLICANT: Recker, Robert R.
; APPLICANT: Johnson, Mark L.
; TITLE OF INVENTION: High bone mass gene of 11q13.3
; FILE REFERENCE: 032796-014
; CURRENT APPLICATION NUMBER: US/09/543,771B
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 09/229,319
; PRIOR FILING DATE: 1999-01-13
; PRIOR APPLICATION NUMBER: US 60/071,449
; PRIOR FILING DATE: 1998-01-13
; PRIOR APPLICATION NUMBER: US 60/105,511
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 641
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1615
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-543-771B-4

Query Match      99.6%; Score 8706; DB 4; Length 1615;
```

Best Local Similarity 99.6%; Pred. No. 0;
Matches 1609; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 MEAAPGPPWLLLLLLLLLLLALCCCPAPAAASPLLLPANRRDVRVLVADAGGKLESTIVS 60
DB 1 MEAAPGPPWLLLLLLLLLLLALCCCPAPAAASPLLLPANRRDVRVLVADAGGKLESTIVS 60
QY 61 GLEDAVAVDFGSKGAVTWDVSEBAIKQTYLNQTAAGVQNVVIGSLVSPDGLACDWGK 120
DB 61 GLEDAVAVDFGSKGAVTWDVSEBAIKQTYLNQTAAGVQNVVIGSLVSPDGLACDWGK 120
QY 121 KLYTDSNTRIEVANLNGTSRKVLFWQDLOQPAIALDPAGHYMYTWDGSETPRIERAG 180
DB 121 KLYTDSNTRIEVANLNGTSRKVLFWQDLOQPAIALDPAGHYMYTWDGSETPRIERAG 180
QY 181 MDGSTRKIIIVDSDIYWPNGLTIDLEEQKLYWADAKLSPIHRANLDGSPROKRVGSLTTP 240
DB 181 MDGSTRKIIIVDSDIYWPNGLTIDLEEQKLYWADAKLSPIHRANLDGSPROKRVGSLTTP 240
QY 241 FALTSLGDTLYTWDQTRSIIHACNKRGTGKKEILSALYSPMDIQVLSQERQPFPHTRCE 300
DB 241 FALTSLGDTLYTWDQTRSIIHACNKRGTGKKEILSALYSPMDIQVLSQERQPFPHTRCE 300
QY 301 EDNGGCSHLCLSSEPPYTACPTGVOLQNGRTCKAGABEVLLARTRDLRISLDT 360
DB 301 EDNGGCSHLCLSSEPPYTACPTGVOLQNGRTCKAGABEVLLARTRDLRISLDT 360
QY 361 DFTDILVQVDDIRHAIDYDPLGYYVWTDDEVRATRRAYLDGSGAQLVNTINPDG 420
DB 361 DFTDILVQVDDIRHAIDYDPLGYYVWTDDEVRATRRAYLDGSGAQLVNTINPDG 420
QY 421 IAVDWARNLYTDTGTDRIEVTRNGTSRKILVSDELDEPRALHPVWGLMYTWDGE 480
DB 421 IAVDWARNLYTDTGTDRIEVTRNGTSRKILVSDELDEPRALHPVWGLMYTWDGE 480
QY 481 NPKIECANLDQERRVLVNASLGNPGLALDQEGKLYWDAKTKIEVINVDGTRKRTL 540
DB 481 NPKIECANLDQERRVLVNASLGNPGLALDQEGKLYWDAKTKIEVINVDGTRKRTL 540
QY 541 LEDKLPHEFTGLGDFRYTWDQRRSIRVHKVQASRDVLIIDQLPDLMLKAVNAKV 600
DB 541 LEDKLPHEFTGLGDFRYTWDQRRSIRVHKVQASRDVLIIDQLPDLMLKAVNAKV 600
QY 601 GTNFCADRNGCSHLCPFTPHATRCGCPIGLELLSDMKTCIVPEAFVFTSRAAHRISL 660
DB 601 GTNFCADRNGCSHLCPFTPHATRCGCPIGLELLSDMKTCIVPEAFVFTSRAAHRISL 660
QY 661 ETNNNDVAIPLTGVKEASALDFVSNHHIYWDVSLKTIISAPFVNGSSVEHVFEGLDYP 720
DB 661 ETNNNDVAIPLTGVKEASALDFVSNHHIYWDVSLKTIISAPFVNGSSVEHVFEGLDYP 720
QY 721 EGMADVWNGKLYWADTGTNRIEVARLDGQPROVLVWRDLNPRSLALDPKGYIYTW 780
DB 721 EGMADVWNGKLYWADTGTNRIEVARLDGQPROVLVWRDLNPRSLALDPKGYIYTW 780
QY 781 GKGPRIVRAFMGNTNCTLVKGRANDLTIDYADQLYWTDLTNMIESNNMLGQERV 840
DB 781 GKGPRIVRAFMGNTNCTLVKGRANDLTIDYADQLYWTDLTNMIESNNMLGQERV 840
QY 841 IADLPHFPGLTQYSDIYIYTDWNLHLSIERADKTSGRNRTLIQGLHDFVMDILVFHSSRQ 900
DB 841 IADLPHFPGLTQYSDIYIYTDWNLHLSIERADKTSGRNRTLIQGLHDFVMDILVFHSSRQ 900
QY 901 DGLNDCHNNGCQOLCLAIIPGHRCCGASHYTLDPSSRNCSPPTLLFSQKSAISRM 960
DB 901 DGLNDCHNNGCQOLCLAIIPGHRCCGASHYTLDPSSRNCSPPTLLFSQKSAISRM 960
QY 961 PDDQHSPLILPLHGLRNVKAIIDYDLDKFIYWDGRQNIKRAKDDGTQPPVLTSQGG 1020
DB 961 PDDQHSPLILPLHGLRNVKAIIDYDLDKFIYWDGRQNIKRAKDDGTQPPVLTSQGG 1020
QY 1021 NPDQPHDLSIDIYSRTLFWTCEATNTINVHRLSGEAMGVVLGRDQKRAIVVNAERG 1080

RESULT 7
US-09-060-299-4
; Sequence 4, Application US/09060299
; Patent No. 6545137
; GENERAL INFORMATION:
; APPLICANT: Todd, John A
; APPLICANT: Hess, John W
; APPLICANT: Caskey, Charles T
; APPLICANT: Cox, Roger D
; APPLICANT: Gerhold, David
; APPLICANT: Hammond, Holly
; APPLICANT: Hey, Patricia
; APPLICANT: Kawaguchi, Yoshihiko
; APPLICANT: Merriman, Tony R
; APPLICANT: Metzker, Michael L
; TITLE OF INVENTION: No. 6545137el Receptor
; NUMBER OF SEQUENCES: 455
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon and Vanderhye
; STREET: 1100 No. 6545137th Glebe Road, Eighth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: US
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/060,299
; FILING DATE: 15-APR-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/043,553

FILING DATE: 15-APR-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/048,740

FILING DATE: 05-JUN-1997

ATTORNEY/AGENT INFORMATION:

NAME: B.J.Sadoff

REGISTRATION NUMBER: 36,663

REFERENCE/DOCKET NUMBER: 620-35

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703)816-4091

TELEFAX: (703)816-4100

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 1591 amino acids

TYPE: amino acid

TOPOLOGY: linear

US-09-060-299-4

Query Match 98.5%; Score 8607; DB 4; Length 1591;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1590; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 25 CPAPAAASPLLLFANRRDRLVLDAGGVKLESTIVVSGLEDAADVDFQPSKGAIVYWDVSE 84
DB 1 CPAPAAASPLLLFANRRDRLVLDAGGVKLESTIVVSGLEDAADVDFQPSKGAIVYWDVSE 60
QY 85 EAIKQTYLNQGAQVNVVIVSGLVSPDGLADWVGKLYWTDSETNRIEVANLNGTSRKV 144
DB 61 EAIKQTYLNQGAQVNVVIVSGLVSPDGLADWVGKLYWTDSETNRIEVANLNGTSRKV 120
QY 145 LFWDQDLPRAIALDPAGHYMTDWTGPRIERAGMDGSTRKLIIVDSDIYWPNGLTIDL 204
DB 121 LFWDQDLPRAIALDPAGHYMTDWTGPRIERAGMDGSTRKLIIVDSDIYWPNGLTIDL 180
QY 205 EEQKLYWADAKLSFTHRANLDGSRQKVEGSLTHPFAITLSGDTLYWTDQTRSIHACN 264
DB 181 EEQKLYWADAKLSFTHRANLDGSRQKVEGSLTHPFAITLSGDTLYWTDQTRSIHACN 240
QY 265 KRTGGRKEIISALYSPMDIQVLQOEROPFFHTRCEEDNGGCSHLCLLSPSEPFYTCAP 324
DB 241 KRTGGRKEIISALYSPMDIQVLQOEROPFFHTRCEEDNGGCSHLCLLSPSEPFYTCAP 300
QY 325 TGVQLQNGRTCKGAEEVLLARLTDLRRLSLDTPDFTDVLQVDDIRHAIAIDYDPLE 384
DB 301 TGVQLQNGRTCKGAEEVLLARLTDLRRLSLDTPDFTDVLQVDDIRHAIAIDYDPLE 360
QY 385 GYVYVTDDEVAIRRAYLDGSGAQTLVNTEINDPDGIADVAVARNLYWTDGTDRIEVTR 444
DB 361 GYVYVTDDEVAIRRAYLDGSGAQTLVNTEINDPDGIADVAVARNLYWTDGTDRIEVTR 420
QY 445 LNSTSRKILVSEDLDEPRAIALHPVGMGLMYWTDGENPKIECANLDGQERRVLYNASLIGW 504
DB 421 LNSTSRKILVSEDLDEPRAIALHPVGMGLMYWTDGENPKIECANLDGQERRVLYNASLIGW 480
QY 505 PNGLALDLQEGKLYWGDAKTDKIEVINVDGTRKRTLLDKLPHIFGFTLLGDFIYWTDMQ 564
DB 481 PNGLALDLQEGKLYWGDAKTDKIEVINVDGTRKRTLLDKLPHIFGFTLLGDFIYWTDMQ 540
QY 565 RRSIERVHKVAKSRDVIIDQLPDLMLGLKANVAVKVGTPNCPADRNGGCSHLCFPTPHATR 624
DB 541 RRSIERVHKVAKSRDVIIDQLPDLMLGLKANVAVKVGTPNCPADRNGGCSHLCFPTPHATR 600
QY 625 CGCPIGLELSDMTKTCIVPEAFVFTSRAAIIHRSLETNNDVAIPITGVKEASALDFDV 684
DB 601 CGCPIGLELSDMTKTCIVPEAFVFTSRAAIIHRSLETNNDVAIPITGVKEASALDFDV 660
QY 685 SNNHIYWDVSLKTIISRAFMNGSSVHVVERGLDYPEGMAVDMGKLYWADTGTNRIEV 744
DB 661 SNNHIYWDVSLKTIISRAFMNGSSVHVVERGLDYPEGMAVDMGKLYWADTGTNRIEV 720
QY 745 ARLDGQFQVLYWDLNPNRSLALDPTKGYIYWTWEGGKPRIVAFMDGTNCMTLVDKVG 804
DB 721 ARLDGQFQVLYWDLNPNRSLALDPTKGYIYWTWEGGKPRIVAFMDGTNCMTLVDKVG 780

QY 805 RANDLTIDYADQRLYWTDLDTNMISSNMLGQERVIADDLPHFPGLTQYSDIYIYWTWN 864
DB 781 RANDLTIDYADQRLYWTDLDTNMISSNMLGQERVIADDLPHFPGLTQYSDIYIYWTWN 840
QY 865 LHSIERADKTSGRNRTLQGLHDFVMDILVPHSSRODGLNDCMNNNGCGQLCLAIPOGH 924
DB 841 LHSIERADKTSGRNRTLQGLHDFVMDILVPHSSRODGLNDCMNNNGCGQLCLAIPOGH 900
QY 925 RCGCASHYTLDPSSRNCSPPFTTLLFSOKSAISMIIPDOQHSPLILPLHGRNVKADY 984
DB 901 RCGCASHYTLDPSSRNCSPPFTTLLFSOKSAISMIIPDOQHSPLILPLHGRNVKADY 960
QY 985 DPLDKFIYWDGRQNIKRAKDDGTQPFVLTSLSQGQMPDRQPHDLSIDIYRTLFWTCEA 1044
DB 961 DPLDKFIYWDGRQNIKRAKDDGTQPFVLTSLSQGQMPDRQPHDLSIDIYRTLFWTCEA 1020
QY 1045 TINTINVHRLSGEAMGVLRGDRDKPRAIVVNAERGILYFTNMQDRAAKIERAALDGTRE 1104
DB 1021 TINTINVHRLSGEAMGVLRGDRDKPRAIVVNAERGILYFTNMQDRAAKIERAALDGTRE 1080
QY 1105 VLFTTGLIRPVALVVDNTLGLKFWVDADLKRIESCDLSGANRLTLEDANIYQPLGLTILG 1164
DB 1081 VLFTTGLIRPVALVVDNTLGLKFWVDADLKRIESCDLSGANRLTLEDANIYQPLGLTILG 1140
QY 1165 KHLWIDRQOQMIERVEKTTGDKRTRIQRVAHLTGIHAEVEVSELEFSAPHCARDNGGC 1224
DB 1141 KHLWIDRQOQMIERVEKTTGDKRTRIQRVAHLTGIHAEVEVSELEFSAPHCARDNGGC 1200
QY 1225 SHICIAKGDGTPRCSCPVLVLLQNLTCGEPPTCSPOFACATGEIDCIPGAWRCDFP 1284
DB 1201 SHICIAKGDGTPRCSCPVLVLLQNLTCGEPPTCSPOFACATGEIDCIPGAWRCDFP 1260
QY 1285 ECDQSDDEEGCPVCSAAQFPCCARGQCVDLRLRCDEADCCDRSDEADCDALCLPNQFRCA 1344
DB 1261 ECDQSDDEEGCPVCSAAQFPCCARGQCVDLRLRCDEADCCDRSDEADCDALCLPNQFRCA 1320
QY 1345 SGQCVLIIKQCCDPPDCIDGSDMLMCEITKPPSDSDSAHSAISGPIVIGIILSLFVMSGVY 1404
DB 1321 SGQCVLIIKQCCDPPDCIDGSDMLMCEITKPPSDSDSAHSAISGPIVIGIILSLFVMSGVY 1380
QY 1405 FVQORVVCQRYAGANGFPFHEYVSGTPHVPINFTAPGSGOHGPTGTACGKSMSSVSLM 1464
DB 1381 FVQORVVCQRYAGANGFPFHEYVSGTPHVPINFTAPGSGOHGPTGTACGKSMSSVSLM 1440
QY 1465 GGRGGVLYDRNHVHTGASSSSSSSTKATLYPILNPPSPATDPSLYNMDMFYSSNIPAT 1524
DB 1441 GGRGGVLYDRNHVHTGASSSSSSSTKATLYPILNPPSPATDPSLYNMDMFYSSNIPAT 1500
QY 1525 ARPYRPIIIRGMAPPPTPCSTIDVCDSDYSASRWKASKYIYDLNDSDDYPYPPPPPHSOYL 1584
DB 1501 VRPYRPIIIRGMAPPPTPCSTIDVCDSDYSASRWKASKYIYDLNDSDDYPYPPPPPHSOYL 1560
QY 1585 SAEDSCPPSPATERSYFHLFPFPPSPCTDSS 1615
DB 1561 SAEDSCPPSPATERSYFHLFPFPPSPCTDSS 1591

RESULT 8

US-09-060-299-43

; Sequence 43, Application US/09060299

; Patent No. 6545137

; GENERAL INFORMATION:

; APPLICANT: Todd, John A

; APPLICANT: Hess, John W

; APPLICANT: Caskey, Charles T

; APPLICANT: Cox, Roger D

; APPLICANT: Gerhold, David

; APPLICANT: Hammond, Holly

; APPLICANT: Hey, Patricia

; APPLICANT: Kawaguchi, Yoshihiko

; APPLICANT: Merriman, Tony R

; APPLICANT: Metzker, Michael L

TITLE OF INVENTION: No. 6545137e1 Receptor
NUMBER OF SEQUENCES: 455
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon and Vanderhye
STREET: 1100 No. 6545137th Glebe Road, Eighth Floor
CITY: Arlington
STATE: Virginia
COUNTRY: US
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/060,299
FILING DATE: 15-APR-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/043,553
FILING DATE: 15-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/048,740
FILING DATE: 05-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: B.J.Sadoff
REGISTRATION NUMBER: 36,663
REFERENCE/DOCKET NUMBER: 620-35
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)816-4091
TELEFAX: (703)816-4100
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 1591 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-09-060-299-43

Query Match 98.5%; Score 8607; DB 4; Length 1591;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1590; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 25 CPAPAAASPLLLFANRRDRLVLDAGGVKLESTIVVSGLEDAADVDFQFSKGAIVYTDVSE 84
Db 1 CPAPAAASPLLLFANRRDRLVLDAGGVKLESTIVVSGLEDAADVDFQFSKGAIVYTDVSE 60
Qy 85 EAIKQTYLNQTTGAQVNVISGLVSPDGLACDWYVGLVYTDSTNRIEVAANLNGTSRKV 144
Db 61 EAIKQTYLNQTTGAQVNVISGLVSPDGLACDWYVGLVYTDSTNRIEVAANLNGTSRKV 120
Qy 145 LFWDQLDQPRALDPAHGYWYTDWGETPRIERAGMDGSTRKIIIVDSIIYVWPNGLTIDL 204
Db 121 LFWDQLDQPRALDPAHGYWYTDWGETPRIERAGMDGSTRKIIIVDSIIYVWPNGLTIDL 180
Qy 205 EEOKLYWADAKLSFIHRANLDGSPQKVEGSLTHFPFALTLSGDTLYWTDWQTSIHACN 264
Db 181 EEOKLYWADAKLSFIHRANLDGSPQKVEGSLTHFPFALTLSGDTLYWTDWQTSIHACN 240
Qy 265 KRTGKKEKILSALYSPMDIQVLSQERQPPFHTRCEDNGGCSHCLLSPEPPTCACP 324
Db 241 KRTGKKEKILSALYSPMDIQVLSQERQPPFHTRCEDNGGCSHCLLSPEPPTCACP 300
Qy 325 TGVQLQNGRTCKAGAEVLLARRTDLRRISLDTDPDTDLVLOVDDIRHAIADYDPLE 384
Db 301 TGVQLQNGRTCKAGAEVLLARRTDLRRISLDTDPDTDLVLOVDDIRHAIADYDPLE 360
Qy 385 GYVYVTDDEVAIRRAVLDGSGAQTLYNTEINDPDGIADVWVARNLYWTDGTDRIEVTR 444
Db 361 GYVYVTDDEVAIRRAVLDGSGAQTLYNTEINDPDGIADVWVARNLYWTDGTDRIEVTR 420
Qy 445 LNTGTSRKILVSEDLDEPRALIAHPVWMLYWDGNEPKIECANLDGQERRVLVNASLIGW 504
Db 421 LNTGTSRKILVSEDLDEPRALIAHPVWMLYWDGNEPKIECANLDGQERRVLVNASLIGW 480

Qy 505 PNGIALDLQEGKLYWGDAKTDKIEVINVDGTRKRTLLDKLPHIFGFTLLGDFIYTDWQ 564
Db 481 PNGIALDLQEGKLYWGDAKTDKIEVINVDGTRKRTLLDKLPHIFGFTLLGDFIYTDWQ 540
Qy 565 RRSIERVHKVKSARDVIIDQLPDLMLGLKAVNVAVKVGTPNCPADRNGGCSHLCFTTPHATR 624
Db 541 RRSIERVHKVKSARDVIIDQLPDLMLGLKAVNVAVKVGTPNCPADRNGGCSHLCFTTPHATR 600
Qy 625 CGCPIGLELLSDMKTCIVPRAFLVFTSRAAIIHRSLETTNNNDVAIPLTGKVEASALDFDV 684
Db 601 CGCPIGLELLSDMKTCIVPRAFLVFTSRAAIIHRSLETTNNNDVAIPLTGKVEASALDFDV 660
Qy 685 SNNHIIYTDVSLKTIISRAFMNGSSVEHVVEFGLDYPGMAVDMGKNLYWADGTGNRIEV 744
Db 661 SNNHIIYTDVSLKTIISRAFMNGSSVEHVVEFGLDYPGMAVDMGKNLYWADGTGNRIEV 720
Qy 745 ARLDGQPRQVLVWRDLNPRSLALDPTKGYIYTWENGKPKRIVRAFMDGTNCMTLVKVG 804
Db 721 ARLDGQPRQVLVWRDLNPRSLALDPTKGYIYTWENGKPKRIVRAFMDGTNCMTLVKVG 780
Qy 805 RANDLTIDYADORLYWTDLTNMISSNMLGQBRVVIADDLPHFPGLTOYSDIYTWTDN 864
Db 781 RANDLTIDYADORLYWTDLTNMISSNMLGQBRVVIADDLPHFPGLTOYSDIYTWTDN 840
Qy 865 LHSIERADKTSGRNRTLIQGHLDVMDILVFHSSRQDGLNDCMENNCGCQGLCLAIPEGH 924
Db 841 LHSIERADKTSGRNRTLIQGHLDVMDILVFHSSRQDGLNDCMENNCGCQGLCLAIPEGH 900
Qy 925 RGCASHYTLDPSSRNCSPPTFLFPQKSAISMIIPDDQHSPLDILPLHGLRNVAIDY 984
Db 901 RGCASHYTLDPSSRNCSPPTFLFPQKSAISMIIPDDQHSPLDILPLHGLRNVAIDY 960
Qy 985 DPLDKFIYWDGRONIKRAKDDGTQPFVLTSLSGQNPDPQPHDLSIDYISRTLPWTCEA 1044
Db 961 DPLDKFIYWDGRONIKRAKDDGTQPFVLTSLSGQNPDPQPHDLSIDYISRTLPWTCEA 1020
Qy 1045 TINTINVHRLSCEAMGVVLRGDRDKPRAIVVNAERGILYFTNMQDRAAKIERAALDGTRE 1104
Db 1021 TINTINVHRLSCEAMGVVLRGDRDKPRAIVVNAERGILYFTNMQDRAAKIERAALDGTRE 1080
Qy 1105 VLFTTGLIRPVALVVDNTLQKLFWVDADLKRIBSCDLSGANRLTLEDANIYVQPLGLTILG 1164
Db 1081 VLFTTGLIRPVALVVDNTLQKLFWVDADLKRIBSCDLSGANRLTLEDANIYVQPLGLTILG 1140
Qy 1165 KHLIYWDROQOMIERVEKTTGDKETRIQGRVAHLTGTHAVEEVELEEFSAHPCARDNGCC 1224
Db 1141 KHLIYWDROQOMIERVEKTTGDKETRIQGRVAHLTGTHAVEEVELEEFSAHPCARDNGCC 1200
Qy 1225 SHICIAKGDGTTPRCPCPVHLVLLQNLLTCGEPPTCSPDQFACATGEIDCIPGAWRCDFGP 1284
Db 1201 SHICIAKGDGTTPRCPCPVHLVLLQNLLTCGEPPTCSPDQFACATGEIDCIPGAWRCDFGP 1260
Qy 1285 ECDQSDDEEGCPVCSAAQFPFCARGCQVDLRLRCDEADCDQSDDEADCDAI CLNPQFRCA 1344
Db 1261 ECDQSDDEEGCPVCSAAQFPFCARGCQVDLRLRCDEADCDQSDDEADCDAI CLNPQFRCA 1320
Qy 1345 SGQCVLTKQCDSPDCIDGSDDELWCEITKPPSDSDSAHSSAIGPVIGIILSLFVWGGVY 1404
Db 1321 SGQCVLTKQCDSPDCIDGSDDELWCEITKPPSDSDSAHSSAIGPVIGIILSLFVWGGVY 1380
Qy 1405 FVQORVVCQYAGANGPFPHEYVSGTTPHVPLNFTAPGSGHQHPTGIACGKSMSSVSLM 1464
Db 1381 FVQORVVCQYAGANGPFPHEYVSGTTPHVPLNFTAPGSGHQHPTGIACGKSMSSVSLM 1440
Qy 1465 GGRGVPLYDRNHVTVGASSSSSTKATLYPPIINLPPSPATDPSLYNMOMFYSSNIPAT 1524
Db 1441 GGRGVPLYDRNHVTVGASSSSSTKATLYPPIINLPPSPATDPSLYNMOMFYSSNIPAT 1500
Qy 1525 ARPYRPVYIIGMAAPPPTPCSTDVCDSDYSASRWKASKYVLDLNSDSDPYPPPPPHSOYL 1584
Db 1501 VRPYRPVYIIGMAAPPPTPCSTDVCDSDYSASRWKASKYVLDLNSDSDPYPPPPPHSOYL 1560
Qy 1585 SAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615

Db 1561 SAEDSCPPSPATERSYFHLFPFPPSPCTDSS 1591

RESULT 9
US-09-402-923A-4
; Sequence 4, Application US/09402923A
; Patent No. 6555654
; GENERAL INFORMATION:
; APPLICANT: Todd, John A
; Hess, John W
; Caskey, Charles T
; Cox, Roger D
; Gerhold, David
; Hammond, Holly
; Hey, Patricia
; Kawaguchi, Yoshihiko
; Merriman, Tony R
; Metzker, Michael L
; TITLE OF INVENTION: No. 6555654e1 LDL-Receptor
; NUMBER OF SEQUENCES: 455
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon and Vanderhye
; STREET: 1100 No. 6555654th Glebe Road, Eighth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: US
; ZIP: VA 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA: US/09/402,923A
; FILING DATE: 14-Feb-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB98/01102
; FILING DATE: 15-APR-1998
; APPLICATION NUMBER: US 60/043,553
; FILING DATE: 15-APR-1997
; APPLICATION NUMBER: US 60/048,740
; FILING DATE: 05-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: B.J.Sadoff
; REGISTRATION NUMBER: 36,663
; REFERENCE/DOCKET NUMBER: 620-81
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)816-4091
; TELEFAX: (703)816-4100
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1591 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-402-923A-4

Query Match 98.5%; Score 8607; DB 4; Length 1591;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1590; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 25 CPAPAAASPLLLFANRRDRLVLDAGGVKLESTIVVSGLEDAAAVDFQFSKGAIVYTDVSE 84
Db 1 CPAPAAASPLLLFANRRDRLVLDAGGVKLESTIVVSGLEDAAAVDFQFSKGAIVYTDVSE 60

Qy 85 EAIQTILNQGAQVNVISGLVSPDGLACDWDVKGKLYWTDSTNRIEVANLNGTSRKV 144
Db 61 EAIQTILNQGAQVNVISGLVSPDGLACDWDVKGKLYWTDSTNRIEVANLNGTSRKV 120

Qy 145 LFWDQDLPRAIALDPAGHYWYWDWGETPRIERAGMDGSTRKIIIVDSDIYWPNGLTIDL 204
Db 121 LFWDQDLPRAIALDPAGHYWYWDWGETPRIERAGMDGSTRKIIIVDSDIYWPNGLTIDL 180

Qy 205 EEQKLYWADAKLSFIHRANLDGSPQKVEGSLTHPFAITLSGDTLYWTDQTRSIHACN 264
Db 181 EEQKLYWADAKLSFIHRANLDGSPQKVEGSLTHPFAITLSGDTLYWTDQTRSIHACN 240

Qy 265 KRTGGRKEIISALYSPMDIOVLQSEROPFPHTRCEDNGGCSHCLLSPEBPFTYCAP 324
Db 241 KRTGGRKEIISALYSPMDIOVLQSEROPFPHTRCEDNGGCSHCLLSPEBPFTYCAP 300

Qy 325 TGVQLQDNQRTCKAGAEVLLARRTDLRRISLQTPDFTDVLQVDDIRHAIADYDPLE 384
Db 301 TGVQLQDNQRTCKAGAEVLLARRTDLRRISLQTPDFTDVLQVDDIRHAIADYDPLE 360

Qy 385 GYVYWTDDVRAIRAYLDGSGAQTLLVNTENDPGIAVDMVARNLYWTDGTDRIEVTR 444
Db 361 GYVYWTDDVRAIRAYLDGSGAQTLLVNTENDPGIAVDMVARNLYWTDGTDRIEVTR 420

Qy 445 LNSTSRKILVSEDLDEPAIAHPVMGLMYWTDGENPKIECANLDGQERRVIVNASLGM 504
Db 421 LNSTSRKILVSEDLDEPAIAHPVMGLMYWTDGENPKIECANLDGQERRVIVNASLGM 480

Qy 505 PNGLALDLOEGKLYWDAKTDKIEVINVDGTRKRTLLKPLHIFGFTLLGDFIYWTQW 564
Db 481 PNGLALDLOEGKLYWDAKTDKIEVINVDGTRKRTLLKPLHIFGFTLLGDFIYWTQW 540

Qy 565 RRSIERVHKVKAASRDVIIDQLPDLMLGLKAVNAVAVGTNPACDRNGGCSHLCFTTPHATR 624
Db 541 RRSIERVHKVKAASRDVIIDQLPDLMLGLKAVNAVAVGTNPACDRNGGCSHLCFTTPHATR 600

Qy 625 CGCFIGLELLSDMKTCIVPEAFVFTSRAAIIHRSLETNNNDVAIPLTGKVEASALDFDV 684
Db 601 CGCFIGLELLSDMKTCIVPEAFVFTSRAAIIHRSLETNNNDVAIPLTGKVEASALDFDV 660

Qy 685 SNHHIYWTDVSLKTIISRAFNMGSSVEHVVEFGLDYPEGMAVDMGKLYWADTGNRIEV 744
Db 661 SNHHIYWTDVSLKTIISRAFNMGSSVEHVVEFGLDYPEGMAVDMGKLYWADTGNRIEV 720

Qy 745 ARLDGQFQVILVWRDLNPRSLALDPTKGYIYTWENGCKPRIVAFMDGTNCMTLVKVG 804
Db 721 ARLDGQFQVILVWRDLNPRSLALDPTKGYIYTWENGCKPRIVAFMDGTNCMTLVKVG 780

Qy 805 RANDLTIDYADQRLYWTDLDTNMISSNMLGQERVIADLLPHFPGLTQYSDIYWTQW 864
Db 781 RANDLTIDYADQRLYWTDLDTNMISSNMLGQERVIADLLPHFPGLTQYSDIYWTQW 840

Qy 865 LHSIERADKTSGRNRTLIQGHLPFVMDILVFHSSRQDGLNDCMHNNGCCQCLAIPEGH 924
Db 841 LHSIERADKTSGRNRTLIQGHLPFVMDILVFHSSRQDGLNDCMHNNGCCQCLAIPEGH 900

Qy 925 RCGCASHYTLDPSSRNCSPPTTFLFSQSAISRMIPDDOHSPLIILPLHGLRNVKAI 984
Db 901 RCGCASHYTLDPSSRNCSPPTTFLFSQSAISRMIPDDOHSPLIILPLHGLRNVKAI 960

Qy 985 DPLDKFIYVVDGRQNIKRAKDDGTQPFVLTLSQGNPDRQPHDLSIDIYSRTLFWTCEA 1044
Db 961 DPLDKFIYVVDGRQNIKRAKDDGTQPFVLTLSQGNPDRQPHDLSIDIYSRTLFWTCEA 1020

Qy 1045 TTNINVRHLSGEAMGVVLRGDRDKPRAIVVNAERGILYFTNMQDRAKIERAALDGTRE 1104
Db 1021 TTNINVRHLSGEAMGVVLRGDRDKPRAIVVNAERGILYFTNMQDRAKIERAALDGTRE 1080

Qy 1105 VLFTTGLIRPVALLVNDNTLGLFWVDADLKRISCDLSGANRLTLEDANTVQPLGLTILG 1164
Db 1081 VLFTTGLIRPVALLVNDNTLGLFWVDADLKRISCDLSGANRLTLEDANTVQPLGLTILG 1140

Qy 1165 KHLIYWDROQMIERVEKTTGDKRTRIQGRVAHLTGTHAVEVSLEEFSAHPCARDNGGC 1224
Db 1141 KHLIYWDROQMIERVEKTTGDKRTRIQGRVAHLTGTHAVEVSLEEFSAHPCARDNGGC 1200

Qy 1225 SHICIAKGDTFPCSCPVHLVLIQNLITCGEPTCPDQFACATGETDCITPGAWRCDFGP 1284
Db 1201 SHICIAKGDTFPCSCPVHLVLIQNLITCGEPTCPDQFACATGETDCITPGAWRCDFGP 1260

Qy 1285 ECDQSDSEGCFCVCSAAFFPCARGCQVDLRLRCDEADCQDRSDEADCDAICLNPQFCA 1344
Db 1261 ECDQSDSEGCFCVCSAAFFPCARGCQVDLRLRCDEADCQDRSDEADCDAICLNPQFCA 1320
Qy 1345 SGQCVLKQCCSDPDCIDGSDLMCEITKPPSDSDSPAHSSAIGPVIGIILSLFVMGGVY 1404
Db 1321 SGQCVLKQCCSDPDCIDGSDLMCEITKPPSDSDSPAHSSAIGPVIGIILSLFVMGGVY 1380
Qy 1405 FVQCORVQRYAGANGPPPHYVSTPHVPLNFTAPGSGHQGPFTGTACGKSMSSVSLM 1464
Db 1381 FVQCORVQRYAGANGPPPHYVSTPHVPLNFTAPGSGHQGPFTGTACGKSMSSVSLM 1440
Qy 1465 GGRGVPLYDRNHVTVGASSSSSSSTKATLYPILNPPSPATDPSLYNMDFYSSNIPAT 1524
Db 1441 GGRGVPLYDRNHVTVGASSSSSSSTKATLYPILNPPSPATDPSLYNMDFYSSNIPAT 1500
Qy 1525 ARPYRPIYIRGMAPPPTPCSTDVCDSDYSASRWKASKYLLDLNSDSDPYPPPTPHSQYL 1584
Db 1501 VRPYRPIYIRGMAPPPTPCSTDVCDSDYSASRWKASKYLLDLNSDSDPYPPPTPHSQYL 1560
Qy 1585 SAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615
Db 1561 SAEDSCPPSPATERSYFHLPPPPSPCTDSS 1591

RESULT 10

US-09-402-923A-43
; Sequence 43, Application US/09402923A
; Patent No. 655654
; GENERAL INFORMATION:
; APPLICANT: Todd, John A
; Hess, John W
; Caskey, Charles T
; Cox, Roger D
; Gerhold, David
; Hammond, Holly
; Hey, Patricia
; Kawaguchi, Yoshihiko
; Merriman, Tony R
; Metzker, Michael L
; TITLE OF INVENTION: No. 6555654e1 LDL-Receptor
; NUMBER OF SEQUENCES: 455
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon and Vanderhye
; STREET: 1100 No. 6555654th Glebe Road, Eighth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: US
; ZIP: VA 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/402,923A
; FILING DATE: 14-Feb-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB98/01102
; FILING DATE: 15-APR-1998
; APPLICATION NUMBER: US 60/043,553
; FILING DATE: 15-APR-1997
; APPLICATION NUMBER: US 60/048,740
; FILING DATE: 05-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: B.J.Sadoff
; REGISTRATION NUMBER: 36,663
; REFERENCE/DOCKET NUMBER: 620-81
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)816-4091
; TELEFAX: (703)816-4100
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:

; LENGTH: 1591 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 43:
US-09-402-923A-43

Query Match
Best Local Similarity 98.5%; Score 8607; DB 4; Length 1591;
Matches 1590; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 25 CPAPAAASPLLLFANRRDRLVDAGGVKLESTIVVSGLEDAADAAVDFQPSKGAIVYWDVSE 84
Db 1 CPAPAAASPLLLFANRRDRLVDAGGVKLESTIVVSGLEDAADAAVDFQPSKGAIVYWDVSE 60
Qy 85 EAIKQTYLNOTGAQVNVISGLVSPDGLACDWGKKLYWTDSETNRLEVANLNGTSRKV 144
Db 61 EAIKQTYLNOTGAQVNVISGLVSPDGLACDWGKKLYWTDSETNRLEVANLNGTSRKV 120
Qy 145 LFMQDLDPQRAIALDPAHGYMYWTDWGETPRIERAGMDGSTRKIIIVDSDIYWPNGLTIDL 204
Db 121 LFMQDLDPQRAIALDPAHGYMYWTDWGETPRIERAGMDGSTRKIIIVDSDIYWPNGLTIDL 180
Qy 205 EEQKLYWADAKLSPIHRANLDGSPQKVVEGSLTHPPFALTLSGDTLYWTDQWTRSIHACN 264
Db 181 EEQKLYWADAKLSPIHRANLDGSPQKVVEGSLTHPPFALTLSGDTLYWTDQWTRSIHACN 240
Qy 265 KRTGGKKEILSALYSAMDIOVLQVLSOERQPPFHTCEEDNGGCSHLCLLSPSEPFYTCACP 324
Db 241 KRTGGKKEILSALYSAMDIOVLQVLSOERQPPFHTCEEDNGGCSHLCLLSPSEPFYTCACP 300
Qy 325 TGVQLQNGRTCKAGAEVLLARRTDLRLISLTPDFTDIVLQVDDIRHAIADYDPLE 384
Db 301 TGVQLQNGRTCKAGAEVLLARRTDLRLISLTPDFTDIVLQVDDIRHAIADYDPLE 360
Qy 385 GYVYWTDDVRAIRRAYLDGSGAQLVNTENIDPDGIAVDVARNLYWTDGTDRIVETR 444
Db 361 GYVYWTDDVRAIRRAYLDGSGAQLVNTENIDPDGIAVDVARNLYWTDGTDRIVETR 420
Qy 445 LNGTSRKILVSEDLDEPRAIALHPVMGLMYWTDGENPKIECANLDGGERVVLNVLNAGW 504
Db 421 LNGTSRKILVSEDLDEPRAIALHPVMGLMYWTDGENPKIECANLDGGERVVLNVLNAGW 480
Qy 505 PNGLALDLQEGKLYWGDAKTKIEVINVDGTRKRTLLEDKLPKIFGFTLLGDFIYWTDMQ 564
Db 481 PNGLALDLQEGKLYWGDAKTKIEVINVDGTRKRTLLEDKLPKIFGFTLLGDFIYWTDMQ 540
Qy 565 RRSIERVHKVKSARDVTDQLPDLMLGLKAVNVAKVGTNPCCADNRNGGCSHLCPFTPHATR 624
Db 541 RRSIERVHKVKSARDVTDQLPDLMLGLKAVNVAKVGTNPCCADNRNGGCSHLCPFTPHATR 600
Qy 625 CGCPIGLELLSDMKTCIVPRAFLVFTSRAA IHRISLETNNNDVAI PLTGKVEASALDFDV 684
Db 601 CGCPIGLELLSDMKTCIVPRAFLVFTSRAA IHRISLETNNNDVAI PLTGKVEASALDFDV 660
Qy 685 SNNHYYWTDVSLKTIISRAFMNGSSVEHVVEFGLDYPEGMAVDMWGNKLYWADGTGNRIEV 744
Db 661 SNNHYYWTDVSLKTIISRAFMNGSSVEHVVEFGLDYPEGMAVDMWGNKLYWADGTGNRIEV 720
Qy 745 ARLDGQPRQVLVWRDLNPRSLALDPTKGYIYWTWEGGKPRIVRAFMDGTNMTLVDKVG 804
Db 721 ARLDGQPRQVLVWRDLNPRSLALDPTKGYIYWTWEGGKPRIVRAFMDGTNMTLVDKVG 780
Qy 805 RANDLTIDYADQRLYWTDLTNMIESNNMLGOERVJADDLPHFPGLTOYSDIYIYWTOWN 864
Db 781 RANDLTIDYADQRLYWTDLTNMIESNNMLGOERVJADDLPHFPGLTOYSDIYIYWTOWN 840
Qy 865 LHSIERADKTSGRNRTLIQGHLDVFMVILVFHSSRQDGLNDCMNNNGCGQLCLAIPEGH 924
Db 841 LHSIERADKTSGRNRTLIQGHLDVFMVILVFHSSRQDGLNDCMNNNGCGQLCLAIPEGH 900
Qy 925 RCGCASHYTLDPSSRNCSPTTFLFQSKAISRMIPDDQHSPLDILPLHLGRNVKAIIDY 984
Db 901 RCGCASHYTLDPSSRNCSPTTFLFQSKAISRMIPDDQHSPLDILPLHLGRNVKAIIDY 960

QY 985 DPLDKFIYWDGRONIKRAKDDGTQPFVLTSLSQGNPDPROPHDLSIDIYSRTLFWTCBA 1044
DB 961 DPLDKFIYWDGRONIKRAKDDGTQPFVLTSLSQGNPDPROPHDLSIDIYSRTLFWTCBA 1020
QY 1045 TINTINVRHLSGEAMGVVLRGDRDKPRAIVVNAERGVLFTNMDRAAKTIERAALDGTRE 1104
DB 1021 TINTINVRHLSGEAMGVVLRGDRDKPRAIVVNAERGVLFTNMDRAAKTIERAALDGTRE 1080
QY 1105 VLFTTGLIRPVALVONTLGLKFWADALRIESCDLSGANRLTLEDANIVQPLGLTILG 1164
DB 1081 VLFTTGLIRPVALVONTLGLKFWADALRIESCDLSGANRLTLEDANIVQPLGLTILG 1140
QY 1165 KHLNWDROQOMIERVEKTTGDKRTIQRVAHLTGIIHAVEEVSLEEPSAHCARDNGGC 1224
DB 1141 KHLNWDROQOMIERVEKTTGDKRTIQRVAHLTGIIHAVEEVSLEEPSAHCARDNGGC 1200
QY 1225 SHICIAKGGTGPCSVHLVLLQNLITCGEPPTCSPDQFACATGEIDCIPGAWCDGFP 1284
DB 1201 SHICIAKGGTGPCSVHLVLLQNLITCGEPPTCSPDQFACATGEIDCIPGAWCDGFP 1260
QY 1285 ECDQSDREGCVCSAAOPPCARGQCVDLRLRCDEADCDQSDRDEADCDALCLPNQFRCA 1344
DB 1261 ECDQSDREGCVCSAAOPPCARGQCVDLRLRCDEADCDQSDRDEADCDALCLPNQFRCA 1320
QY 1345 SQQCVLIIKQOCDSFPDCIDGSDLMCEITKPSDPSAHSAGIPVIGIILSLFVMGGVY 1404
DB 1321 SQQCVLIIKQOCDSFPDCIDGSDLMCEITKPSDPSAHSAGIPVIGIILSLFVMGGVY 1380
QY 1405 FVQVRVQRYAGANGPPFHEHYSGTTPHVPNFIAFGGSGHQGFTGACGKSMSSVSIM 1464
DB 1381 FVQVRVQRYAGANGPPFHEHYSGTTPHVPNFIAFGGSGHQGFTGACGKSMSSVSIM 1440
QY 1465 GGRGVPLVYDRNHVTGASSSSSSSTKATLYPILNPPSPATDPSLYNMDMFYSSNIPAT 1524
DB 1441 GGRGVPLVYDRNHVTGASSSSSSSTKATLYPILNPPSPATDPSLYNMDMFYSSNIPAT 1500
QY 1525 ARPYRPIIRGMAPPTPCSTDCVSDYSASRWKASKYLLDLNSDSDYPPPPPHSOYL 1584
DB 1501 VRYRPIIRGMAPPTPCSTDCVSDYSASRWKASKYLLDLNSDSDYPPPPPHSOYL 1560
QY 1585 SAEDSCPPSPATERSYFHLFPPPPSPCTDSS 1615
DB 1561 SAEDSCPPSPATERSYFHLFPPPPSPCTDSS 1591

RESULT 11

US-09-060-299-29

Sequence 29, Application US/09060299

Patent No. 6545137

GENERAL INFORMATION:

APPLICANT: Todd, John A

APPLICANT: Hess, John W

APPLICANT: Caskey, Charles T

APPLICANT: Cox, Roger D

APPLICANT: Gerhold, David

APPLICANT: Hammond, Holly

APPLICANT: Hey, Patricia

APPLICANT: Kawaguchi, Yoshihiko

APPLICANT: Merriman, Tony R

APPLICANT: Metzker, Michael L

TITLE OF INVENTION: No. 6545137e1 Receptor

NUMBER OF SEQUENCES: 455

CORRESPONDENCE ADDRESS:

ADDRESSEE: Nixon and Vanderhye

STREET: 1100 No. 6545137th Glebe Road, Eighth Floor

CITY: Arlington

STATE: Virginia

COUNTRY: US

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/060,299

FILING DATE: 15-APR-1998

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/043,553

FILING DATE: 15-APR-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/048,740

FILING DATE: 05-JUN-1997

ATTORNEY/AGENT INFORMATION:

NAME: B.J. Sadoff

REGISTRATION NUMBER: 36,663

REFERENCE/DOCKET NUMBER: 620-35

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 816-4091

TELEFAX: (703) 816-4100

INFORMATION FOR SEQ ID NO: 29:

SEQUENCE CHARACTERISTICS:

LENGTH: 1639 amino acids

TYPE: amino acid

TOPOLOGY: linear

US-09-060-299-29

Query Match 98.1%; Score 8574; DB 4; Length 1639;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 1585; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 26 PAAAAAPLLLPANRRDRLVLDAGVGKLESTIVVSGLEDAADVDFQFSKGAVYMTDVEE 85

DB 50 PSLHNSPPLLFPANRRDRLVLDAGVGKLESTIVVSGLEDAADVDFQFSKGAVYMTDVEE 109

QY 86 AIKQTYLNTQGAAYQNVVIGSLVSPDGLACDWGKLYTWTSETNRIEVANLNGTSRKL 145

DB 110 AIKQTYLNTQGAAYQNVVIGSLVSPDGLACDWGKLYTWTSETNRIEVANLNGTSRKL 169

QY 146 FWQDLQDPRALDPAGHYMYTWDGTPRIERAGMGSTRIIIVSDIYFNGLTIDLE 205

DB 170 FWQDLQDPRALDPAGHYMYTWDGTPRIERAGMGSTRIIIVSDIYFNGLTIDLE 229

QY 206 EQKLYWADAKLSFTHRANLDGSPQKVVESGLTHPFFALTLSGDTLYWTDWTRIHCNK 265

DB 230 EQKLYWADAKLSFTHRANLDGSPQKVVESGLTHPFFALTLSGDTLYWTDWTRIHCNK 289

QY 266 RTGGRKEILLSALYSPMDIQVLSQERQPFPHTRCEEEDNGGCSHLCLLSPSPFFTCAPT 325

DB 290 RTGGRKEILLSALYSPMDIQVLSQERQPFPHTRCEEEDNGGCSHLCLLSPSPFFTCAPT 349

QY 326 GVQLQDNGRTCKAGAEVLLIARLTDLRRISLDPDFTDIVLQVDDIRHAIADYDPLEG 385

DB 350 GVQLQDNGRTCKAGAEVLLIARLTDLRRISLDPDFTDIVLQVDDIRHAIADYDPLEG 409

QY 386 YVYWTDDVRAIRRAYLDGSGAQTIVNTEINDPGIADVWVARLYWTDGTDRIEVTRL 445

DB 410 YVYWTDDVRAIRRAYLDGSGAQTIVNTEINDPGIADVWVARLYWTDGTDRIEVTRL 469

QY 446 NGTSRKILVSEDLDDEPPAIALHPVMGLMYTWDNGENPKIECANLDGERRVNVNASLWGP 505

DB 470 NGTSRKILVSEDLDDEPPAIALHPVMGLMYTWDNGENPKIECANLDGERRVNVNASLWGP 529

QY 506 NGLALDIQEGKLYWGDATDKIEVINVDGTPKRTLLDKLPHIFGFTLLGDFIYWTWQOR 565

DB 530 NGLALDIQEGKLYWGDATDKIEVINVDGTPKRTLLDKLPHIFGFTLLGDFIYWTWQOR 589

QY 566 RSIERHVKVKSARDVIIDQLPDLMLGKAVNVAKVVGTPNCPADRNGGCSHLCFFTPHATRC 625

DB 590 RSIERHVKVKSARDVIIDQLPDLMLGKAVNVAKVVGTPNCPADRNGGCSHLCFFTPHATRC 649

QY 626 GCPTGLELLSDMKTCTIVPEAFIVFTSRAATHRISLENNNDVAIPLTGVKEASALDSDVS 685

DB 650 GCPTGLELLSDMKTCTIVPEAFIVFTSRAATHRISLENNNDVAIPLTGVKEASALDSDVS 709

QY 386 YVYWTDDVRAIRRAYLDGSGAQLVNTINDEPGIAVDWVARNLYWTGTDRIEVTRL 445
DB 410 YVYWTDDVRAIRRAYLDGSGAQLVNTINDEPGIAVDWVARNLYWTGTDRIEVTRL 469
QY 446 NGTSRKILVSDLEBPRAIALHPVWGLMYTDWGENPKIECANLDGQERRVNVNASLQWP 505
DB 470 NGTSRKILVSDLEBPRAIALHPVWGLMYTDWGENPKIECANLDGQERRVNVNASLQWP 529
QY 506 NGLALDLQEGKLYWGDAKTKEIVNDVGTGRTTLLDKLPHIFGFTLLGDFIYWTWQOR 565
DB 530 NGLALDLQEGKLYWGDAKTKEIVNDVGTGRTTLLDKLPHIFGFTLLGDFIYWTWQOR 589
QY 566 RSIERVHKVKSARDVVIDQLPDLMLGLKAVNVKAVVGTNPNPCADNRGGCSHLCTFFPHATRC 625
DB 590 RSIERVHKVKSARDVVIDQLPDLMLGLKAVNVKAVVGTNPNPCADNRGGCSHLCTFFPHATRC 649
QY 626 GCPGLELLSDMKTCIVPEAFVFTSRAAIIHRSLETNNNDVAIPLTGVKEASALDFDVS 685
DB 650 GCPGLELLSDMKTCIVPEAFVFTSRAAIIHRSLETNNNDVAIPLTGVKEASALDFDVS 709
QY 686 NNHIYWTDVSLKTSRAFNGSSVEHVVEFGLDYPEGMAVDMGKNLYWADTGNRIEVA 745
DB 710 NNHIYWTDVSLKTSRAFNGSSVEHVVEFGLDYPEGMAVDMGKNLYWADTGNRIEVA 769
QY 746 RLDGQFQVLVWRDLNPRSLALDPTKGYIYWTWGGKPRIVRAFMDGTNCTMLVDKVR 805
DB 770 RLDGQFQVLVWRDLNPRSLALDPTKGYIYWTWGGKPRIVRAFMDGTNCTMLVDKVR 829
QY 806 ANDLTIDVADQRLYWTDLTNMIESNNMLGQERVVIAADLPHFGLTGOYSYIYWTWNL 865
DB 830 ANDLTIDVADQRLYWTDLTNMIESNNMLGQERVVIAADLPHFGLTGOYSYIYWTWNL 889
QY 866 HSIERADTSGNRNLTIOGHLDVMDILVHSSRDGLNDCHMNNQCGQLCLAIPEGHR 925
DB 890 HSIERADTSGNRNLTIOGHLDVMDILVHSSRDGLNDCHMNNQCGQLCLAIPEGHR 949
QY 926 CCASHYLDPSRNCSPPTFLFSQSAISRMIPDQHSPLILPLHGLRNKVAIDYD 985
DB 950 CCASHYLDPSRNCSPPTFLFSQSAISRMIPDQHSPLILPLHGLRNKVAIDYD 1009
QY 986 PLDKFIYWDGQNKIRAKDGTQFVLTSLSQGNPDQPHDLSDIYSRTLFTWCAT 1045
DB 1010 PLDKFIYWDGQNKIRAKDGTQFVLTSLSQGNPDQPHDLSDIYSRTLFTWCAT 1069
QY 1046 NTINVHRLSGEAMGVVLRGDRDKPRAIVVNAERGVLFTNMDRAAKTERAALDGTREY 1105
DB 1070 NTINVHRLSGEAMGVVLRGDRDKPRAIVVNAERGVLFTNMDRAAKTERAALDGTREY 1129
QY 1106 LFTTGLIRPVAVLVNTLGLFWVDADLKRIBSCDLSGANRLTLEDANIVPLGLTILGK 1165
DB 1130 LFTTGLIRPVAVLVNTLGLFWVDADLKRIBSCDLSGANRLTLEDANIVPLGLTILGK 1189
QY 1166 HLYWIDRQOQMLERVEKTTGDKRTRIQGRVAHLTGTHAVEEVSLEFSAHPCARDNGCS 1225
DB 1190 HLYWIDRQOQMLERVEKTTGDKRTRIQGRVAHLTGTHAVEEVSLEFSAHPCARDNGCS 1249
QY 1226 HICIAKGDTPRCSPVHLVLLQNLITCGEPTCSPDQACATGIDICIPGAWRCDGPPE 1285
DB 1250 HICIAKGDTPRCSPVHLVLLQNLITCGEPTCSPDQACATGIDICIPGAWRCDGPPE 1309
QY 1286 CDDQSEEGCPVCSAAQFFCARGQCVDLRLRCGDGRADCCDRDEADCAICLPNQFRCAS 1345
DB 1310 CDDQSEEGCPVCSAAQFFCARGQCVDLRLRCGDGRADCCDRDEADCAICLPNQFRCAS 1369
QY 1346 GQCVLIIKQCDSPDCIDGSDBLMCEITKPPSDDSPAHSATGPVIGIILSLFVNGGVYF 1405
DB 1370 GQCVLIIKQCDSPDCIDGSDBLMCEITKPPSDDSPAHSATGPVIGIILSLFVNGGVYF 1429
QY 1406 VQORVVQRYAGANGFPFHYEYSGTHPVLNFIAPGSGOHPFTGTACGKSNMSSVSLMG 1465
DB 1430 VQORVVQRYAGANGFPFHYEYSGTHPVLNFIAPGSGOHPFTGTACGKSNMSSVSLMG 1489

QY 1466 GRGGVLYDRNVHTGATSSSSSSSTKATLYPPILNPPSPATDPSLYNMDMFYSSNIPATA 1525
DB 1490 GRGGVLYDRNVHTGATSSSSSSSTKATLYPPILNPPSPATDPSLYNMDMFYSSNIPATA 1549
QY 1526 RPYRPIYIRGMAPPYTPCSTDVCDSDYSASRWKASKYIYLDLNSDSDPYPPPTPHSQYLS 1585
DB 1550 RPYRPIYIRGMAPPYTPCSTDVCDSDYSASRWKASKYIYLDLNSDSDPYPPPTPHSQYLS 1609
QY 1586 AEDSCPPSPATERSYFHLFPPPPSPCTDSS 1615
DB 1610 AEDSCPPSPATERSYFHLFPPPPSPCTDSS 1639

RESULT 13

US-09-060-299-39
; Sequence 39, Application US/09060299
; Patent No. 6545137
; GENERAL INFORMATION:
; APPLICANT: Todd, John A
; APPLICANT: Hesse, John W
; APPLICANT: Caskey, Charles T
; APPLICANT: Cox, Roger D
; APPLICANT: Gerhold, David
; APPLICANT: Hammond, Holly
; APPLICANT: Hey, Patricia
; APPLICANT: Kawaguchi, Yoshihiko
; APPLICANT: Merriman, Tony R
; APPLICANT: Metzker, Michael L
; TITLE OF INVENTION: No. 6545137el Receptor
; NUMBER OF SEQUENCES: 455
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon and Vanderhye
; STREET: 1100 No. 6545137th Glebe Road, Eighth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: US
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/060,299
; FILING DATE: 15-APR-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/043,553
; FILING DATE: 15-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/048,740
; FILING DATE: 05-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: B.J.Sadoff
; REGISTRATION NUMBER: 36,663
; REFERENCE/POCKET NUMBER: 620-35
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)816-4091
; TELEFAX: (703)816-4100
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1584 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-09-060-299-39

Query Match 98.0%; Score 8568; DB 4; Length 1584;

Best Local Similarity 99.9%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 1583; Conservative 0;

QY 32 SPLLLFANRRDRLVLDAGVKGLESTIVVSGLEDAADVDFQFSKGAVYWTVDVSEEAIKQTY 91

DB 1 SPLLLFANRRDRLVLDAGVKGLESTIVVSGLEDAADVDFQFSKGAVYWTVDVSEEAIKQTY 60

QY 92 LNQTGAAYONVYVIGSLVSPDGLACDWGKLYWTDSETNRIEVANLNGTSRKVLFWQDLD 151
DB 61 LNQTGAAYONVYVIGSLVSPDGLACDWGKLYWTDSETNRIEVANLNGTSRKVLFWQDLD 120
QY 152 QPRAIALDPANGYMYWTDWGETPPIERAGMDGSTRKIIIVDSDIYWPNGLTIDLBEOQLYW 211
DB 121 QPRAIALDPANGYMYWTDWGETPPIERAGMDGSTRKIIIVDSDIYWPNGLTIDLBEOQLYW 180
QY 212 ADAKLSFTHRANLDSGFRQKVVEGSLTHPPALTLSGDTLYWTDWOTRSHACNKRKTGGR 271
DB 181 ADAKLSFTHRANLDSGFRQKVVEGSLTHPPALTLSGDTLYWTDWOTRSHACNKRKTGGR 240
QY 272 KEILSALYSPMDI QVLSQEROPFFHTRCEEDNGGCSHLCLLSPSEFFYTCACTGVQLQD 331
DB 241 KEILSALYSPMDI QVLSQEROPFFHTRCEEDNGGCSHLCLLSPSEFFYTCACTGVQLQD 300
QY 332 NGRTCKAGAEVLLARLTDLRRISLDPDFTDVLQVDDIRHAIADIDPLEGVYVWTD 391
DB 301 NGRTCKAGAEVLLARLTDLRRISLDPDFTDVLQVDDIRHAIADIDPLEGVYVWTD 360
QY 392 DEVRAIRAYLDGSGAOTLVNTEINDPGIAVDWVARNLYWTDGTGDRIEVTRNLGTSRK 451
DB 361 DEVRAIRAYLDGSGAOTLVNTEINDPGIAVDWVARNLYWTDGTGDRIEVTRNLGTSRK 420
QY 452 ILVSEDLDEPRAIALHPVWGLMYWTDWGENPKIECANLDGQERRVLYVNASLWPNGLALD 511
DB 421 ILVSEDLDEPRAIALHPVWGLMYWTDWGENPKIECANLDGQERRVLYVNASLWPNGLALD 480
QY 512 LOEGKLYWGDAKTIEVINVDGTRKRTLLBCKLPHIFPGFTLLGDFIYWTWQRRSIRV 571
DB 481 LOEGKLYWGDAKTIEVINVDGTRKRTLLBCKLPHIFPGFTLLGDFIYWTWQRRSIRV 540
QY 572 HKVKSARDVIIIDQLPDLMLGKAVNAVGVNPNCADRNGGCSHLCLFFTPHATRCGCPGL 631
DB 541 HKVKSARDVIIIDQLPDLMLGKAVNAVGVNPNCADRNGGCSHLCLFFTPHATRCGCPGL 600
QY 632 ELLSDMKTICIVPEAFVFTSRAAIIHRISLETNNNDVAIPLTGVKEASALDFDVSNHHIY 691
DB 601 ELLSDMKTICIVPEAFVFTSRAAIIHRISLETNNNDVAIPLTGVKEASALDFDVSNHHIY 660
QY 692 TDVSLKTIISRAFPMGSSVEHVVEFGLDYPEGMAVDWVGKNIYADTGTNRIEVARLDGQF 751
DB 661 TDVSLKTIISRAFPMGSSVEHVVEFGLDYPEGMAVDWVGKNIYADTGTNRIEVARLDGQF 720
QY 752 ROVLVWRDLNPRSLALDPTKGYIYWTWEGGKPRIVRAFMDGTNCMTLVKVGANDLTI 811
DB 721 ROVLVWRDLNPRSLALDPTKGYIYWTWEGGKPRIVRAFMDGTNCMTLVKVGANDLTI 780
QY 812 DYADORLYWTDLTNMISSNMLGQERVVIADDLPHFPGLTQYSDIYIYWTWNLHSIERA 871
DB 781 DYADORLYWTDLTNMISSNMLGQERVVIADDLPHFPGLTQYSDIYIYWTWNLHSIERA 840
QY 872 DKTSGRNRTLQGLHDFWMDILVPHSSRDQGLNOCMHNNGCGQLCLAIPOGHRGCGASH 931
DB 841 DKTSGRNRTLQGLHDFWMDILVPHSSRDQGLNOCMHNNGCGQLCLAIPOGHRGCGASH 900
QY 932 YTLDPSSRNCSPPTTFLFSSOKSAISRMIPDOHSPDLILPLHGLRNKVAIDYDPLDKFI 991
DB 901 YTLDPSSRNCSPPTTFLFSSOKSAISRMIPDOHSPDLILPLHGLRNKVAIDYDPLDKFI 960
QY 992 YWVDGRQNIKRAKDDGTQPPVLTLSQGNPDROPHDLISIDYISRTLFWTCEATNTINHV 1051
DB 961 YWVDGRQNIKRAKDDGTQPPVLTLSQGNPDROPHDLISIDYISRTLFWTCEATNTINHV 1020
QY 1052 RLSGEAMGVVLGRDRKPRAIIVNAERGILYFTWQDRAAKIERAALDGTREVERLFTTGL 1111
DB 1021 RLSGEAMGVVLGRDRKPRAIIVNAERGILYFTWQDRAAKIERAALDGTREVERLFTTGL 1080
QY 1112 IRPVAVVDNTLGLKFWVDADLKEIESCDLSGANRLTLEDANIYQPLGLTTLGLHLYWID 1171
DB 1081 IRPVAVVDNTLGLKFWVDADLKEIESCDLSGANRLTLEDANIYQPLGLTTLGLHLYWID 1140
QY 1172 RQOQMIERVENTGDKRTRIQRVAHLTGTHAVEEVSLEEPSAHPCARDNGGCSHICIAK 1231

DB 1141 RQOQMIERVENTGDKRTRIQRVAHLTGTHAVEEVSLEEPSAHPCARDNGGCSHICIAK 1200
QY 1232 GDGTPRCSFVHLVLLQNLITCGBPPTSCSPQOFACATGEIDICIPGAWRCDFPPECDDQSD 1291
DB 1201 GDGTPRCSFVHLVLLQNLITCGBPPTSCSPQOFACATGEIDICIPGAWRCDFPPECDDQSD 1260
QY 1292 BEGPCVCSAOPPCARGOCVDLRLRCDEADQDRSDADDAICLPNQFRCSAQCVLI 1351
DB 1261 BEGPCVCSAOPPCARGOCVDLRLRCDEADQDRSDADDAICLPNQFRCSAQCVLI 1320
QY 1352 KQOCDSPDCIDGSDDELMEITKPPSDSPAHSAGIVGIGIILSLFVMGGYVFCQV 1411
DB 1321 KQOCDSPDCIDGSDDELMEITKPPSDSPAHSAGIVGIGIILSLFVMGGYVFCQV 1380
QY 1412 CORVAGANGPPHEVYVSGTVPVPLNFIAPGCSQHQPFPGIACGKSMSSVSLMGGRGV 1471
DB 1381 CORVAGANGPPHEVYVSGTVPVPLNFIAPGCSQHQPFPGIACGKSMSSVSLMGGRGV 1440
QY 1472 LYDRNHVTGASSSSSSSTKATLYPILNPPSPATDPSLYNMDMYSSNIPATARPYP 1531
DB 1441 LYDRNHVTGASSSSSSSTKATLYPILNPPSPATDPSLYNMDMYSSNIPATARPYP 1500
QY 1532 IIRGWAPPTTPTCSTDVCDSDYSASRWKASKYLLDNLSDSDPYPPPTPHSQYLSAEDSCP 1591
DB 1501 IIRGWAPPTTPTCSTDVCDSDYSASRWKASKYLLDNLSDSDPYPPPTPHSQYLSAEDSCP 1560
QY 1592 PSPATERSYFHLFPPTSPCTDSS 1615
DB 1561 PSPATERSYFHLFPPTSPCTDSS 1584

RESULT 14

US-09-402-923A-39
; Sequence 39, Application US/09402923A

; Patent No. 6555654

; GENERAL INFORMATION:

APPLICANT: Todd, John A

Hess, John W

Caskey, Charles T

Cox, Roger D

Gerhold, David

Hammond, Holly

Hey, Patricia

Kawaguchi, Yoshihiko

Merriman, Tony R

Metzker, Michael L

TITLE OF INVENTION: No. 6555654e1 LDL-Receptor

NUMBER OF SEQUENCES: 455

CORRESPONDENCE ADDRESS:

ADDRESSEE: Nixon and Vanderhye

STREET: 1100 No. 6555654th Glebe Road, Eighth Floor

CITY: Arlington

STATE: Virginia

COUNTRY: US

ZIP: VA 22201-4714

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/402,923A

FILING DATE: 14-Feb-2001

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/GB98/01102

FILING DATE: 15-APR-1998

APPLICATION NUMBER: US 60/043,553

FILING DATE: 15-APR-1997

APPLICATION NUMBER: US 60/048,740

FILING DATE: 05-JUN-1997

ATTORNEY/AGENT INFORMATION:

NAME: B.J.Sadoff

REGISTRATION NUMBER: 36,663
REFERENCE/DOCKET NUMBER: 620-81
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)816-4091
TELEFAX: (703)816-4100
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 1584 amino acids
TYPE: amino acid
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 39:
US-09-402-923A-39

Query Match 98.0%; Score 8568; DB 4; Length 1584; Best Local Similarity 99.9%; Pred. NO. 0; Matches 1583; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
QY 32	SPLLIFANRRDVRVLVDAGGVKLESTIVVSGLEDAADVDFQSKGAVYWDVSEAIKQTY 91
DB 1	SPLLIFANRRDVRVLVDAGGVKLESTIVVSGLEDAADVDFQSKGAVYWDVSEAIKQTY 60
QY 92	LNQTGAQVNVVIGSLVSPDGLACDWGKLYWTDSETNRIEVANLNGTSRKVLFWQDLD 151
DB 61	LNQTGAQVNVVIGSLVSPDGLACDWGKLYWTDSETNRIEVANLNGTSRKVLFWQDLD 120
QY 152	QPRALDPAHYMYWTDWGETPRIERAGMDGSTRKIIIVDSDIYWPNGLTIDLEBOKLYW 211
DB 121	QPRALDPAHYMYWTDWGETPRIERAGMDGSTRKIIIVDSDIYWPNGLTIDLEBOKLYW 180
QY 212	ADAKLSFTHRANLDGSKVQVVEGSLTHPEALTLSGDTLYWTDWTRSIHACNKTGGR 271
DB 181	ADAKLSFTHRANLDGSKVQVVEGSLTHPEALTLSGDTLYWTDWTRSIHACNKTGGR 240
QY 272	KEILSALYSPMDIQVLSQEROPFFHTRCEEDNGGCSHLCLLSPSPFFYTCACTGVQLQD 331
DB 241	KEILSALYSPMDIQVLSQEROPFFHTRCEEDNGGCSHLCLLSPSPFFYTCACTGVQLQD 300
QY 332	NGRTCKAGAEVLLIARRTDLRIISLDPDFDVIQVDDIRHATAIDYDPLEGVVYWD 391
DB 301	NGRTCKAGAEVLLIARRTDLRIISLDPDFDVIQVDDIRHATAIDYDPLEGVVYWD 360
QY 392	DEVRAIRAYLDGSAQTLVNTIENDPGIADVARNLYWTDGTGRIEVRNLNGTSRK 451
DB 361	DEVRAIRAYLDGSAQTLVNTIENDPGIADVARNLYWTDGTGRIEVRNLNGTSRK 420
QY 452	ILVSEDLDEPRALALHPVMGLMYWTDWGENPKIECANLDGQRRVVLNVLASLGNPLALD 511
DB 421	ILVSEDLDEPRALALHPVMGLMYWTDWGENPKIECANLDGQRRVVLNVLASLGNPLALD 480
QY 512	LOEGKLYWDGAKTDKLEVINVDGTRKRTLLDKLPHIFGFTLLGDFIYWDWQRSIERV 571
DB 481	LOEGKLYWDGAKTDKLEVINVDGTRKRTLLDKLPHIFGFTLLGDFIYWDWQRSIERV 540
QY 572	HKVKASRDVIIIDQLPDLMLKAVNAVAVGTNPNPCADNRNGGCSHLCTFFPHATRCGPIGL 631
DB 541	HKVKASRDVIIIDQLPDLMLKAVNAVAVGTNPNPCADNRNGGCSHLCTFFPHATRCGPIGL 600
QY 632	ELLSDMKTCIVPEAFIVFTSRAAIHRSIETNNNDVAIPLTGKESASALDFDVSNNHIY 691
DB 601	ELLSDMKTCIVPEAFIVFTSRAAIHRSIETNNNDVAIPLTGKESASALDFDVSNNHIY 660
QY 692	TDVSLKTIISRAFNGSSVEHVEFGLDYPEGMAVDMWGNKLYWADTGTNRIEVARLDGOF 751
DB 661	TDVSLKTIISRAFNGSSVEHVEFGLDYPEGMAVDMWGNKLYWADTGTNRIEVARLDGOF 720
QY 752	RQVLVWRDLNPRSLALDPTKGIYWTWGGKPRIIVRAFMDGTNCMTLVKVRANDLTI 811
DB 721	RQVLVWRDLNPRSLALDPTKGIYWTWGGKPRIIVRAFMDGTNCMTLVKVRANDLTI 780
QY 812	DYADQRLYWTDLDTNNIESNMLQORVVIADLPHPGITQYSYDIYWDNWLHSIERA 871
DB 781	DYADQRLYWTDLDTNNIESNMLQORVVIADLPHPGITQYSYDIYWDNWLHSIERA 840

RESULT 15
US-09-060-299-42
; Sequence 42, Application US/09060299
; Patent No. 6545137
; GENERAL INFORMATION:
; APPLICANT: Todd, John A
; APPLICANT: Hess, John W
; APPLICANT: Caskey, Charles T
; APPLICANT: Cox, Roger D
; APPLICANT: Gerhold, David
; APPLICANT: Hammond, Holly
; APPLICANT: Hey, Patricia
; APPLICANT: Kawaguchi, Yoshihiko
; APPLICANT: Merriman, Tony R
; APPLICANT: Metzker, Michael L
; TITLE OF INVENTION: No. 6545137el Receptor
; NUMBER OF SEQUENCES: 455
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon and Vanderhye
; STREET: 1100 No. 6545137th Glebe Road, Eighth Floor

QY 872	DKTSGNRNLTIOGHLDPVMDILVFHSSRODGLNDCMHNNGCQGLCLAIPEGHRCGCASH 931
DB 841	DKTSGNRNLTIOGHLDPVMDILVFHSSRODGLNDCMHNNGCQGLCLAIPEGHRCGCASH 900
QY 932	YTLDPSSRNCSPPTTFLFSQKSAISRMIPDDQHSPLILPLHGLRNVAIDAIDYDPLDKFI 991
DB 901	YTLDPSSRNCSPPTTFLFSQKSAISRMIPDDQHSPLILPLHGLRNVAIDAIDYDPLDKFI 960
QY 992	YWDGRONIKRAKDDGTQPFVLTSLSQGNPDQPHDLSIDYSRTLFWTCEATNTINVH 1051
DB 961	YWDGRONIKRAKDDGTQPFVLTSLSQGNPDQPHDLSIDYSRTLFWTCEATNTINVH 1020
QY 1052	RLSGEAMGVVLGRDRKPRAIVVNAERGILYFTNMODRAAKIERAAALDGTREVLFTTGL 1111
DB 1021	RLSGEAMGVVLGRDRKPRAIVVNAERGILYFTNMODRAAKIERAAALDGTREVLFTTGL 1080
QY 1112	IRPVALVVDNTLGLFWVDADLKRIESCDLSGANRLTLEDANIVQPLGLTILGKHLWID 1171
DB 1081	IRPVALVVDNTLGLFWVDADLKRIESCDLSGANRLTLEDANIVQPLGLTILGKHLWID 1140
QY 1172	ROQMIERVEKTTGDKETRIQGRVAHLTGTHAVEEVSLEEFSAHPCARDNGGCSHICIAK 1231
DB 1141	ROQMIERVEKTTGDKETRIQGRVAHLTGTHAVEEVSLEEFSAHPCARDNGGCSHICIAK 1200
QY 1232	GDGTPRCSVHLVLLQNLLTCGEPPTCSPPQFACATGEIDCIPGAWRCDFPECDQSD 1291
DB 1201	GDGTPRCSVHLVLLQNLLTCGEPPTCSPPQFACATGEIDCIPGAWRCDFPECDQSD 1260
QY 1292	BEGCPVCSAQAQPCARGQCVDLRLRCDEADQDRSDEADCAICLPNQFRCASGQCVLI 1351
DB 1261	BEGCPVCSAQAQPCARGQCVDLRLRCDEADQDRSDEADCAICLPNQFRCASGQCVLI 1320
QY 1352	KQCCDSPPDCIDGDELWCETTKPPSDSPAHSSAIGPVGIIISLFVMGGVYFVQCRVV 1411
DB 1321	KQCCDSPPDCIDGDELWCETTKPPSDSPAHSSAIGPVGIIISLFVMGGVYFVQCRVV 1380
QY 1412	CORVAGANGPPPEHYVSGTHVPLNFIAPGSQHPTGTIACGSKMSSVSLMGGRGVVP 1471
DB 1381	CORVAGANGPPPEHYVSGTHVPLNFIAPGSQHPTGTIACGSKMSSVSLMGGRGVVP 1440
QY 1472	LYDRNHVTGASSSSSSSTKATLYPILNPPSPATDPSLYNMDMFSNIPATVPYRPY 1531
DB 1441	LYDRNHVTGASSSSSSSTKATLYPILNPPSPATDPSLYNMDMFSNIPATVPYRPY 1500
QY 1532	IIRGMWAPPTTCSITDVCSDSYASRWKASKYLYDLNDSDDPYPPPTPHSYLSAEDSCP 1591
DB 1501	IIRGMWAPPTTCSITDVCSDSYASRWKASKYLYDLNDSDDPYPPPTPHSYLSAEDSCP 1560
QY 1592	PSPATERSYFHLFPFPPSPCTDSS 1615
DB 1561	PSPATERSYFHLFPFPPSPCTDSS 1584

CITY: Arlington
STATE: Virginia
COUNTRY: US
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US 09/060,299
FILING DATE: 15-APR-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/043,553
FILING DATE: 15-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/048,740
FILING DATE: 05-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: B.J.Sadoff
REGISTRATION NUMBER: 36,663
REFERENCE/DOCKET NUMBER: 620-35
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)816-4091
TELEFAX: (703)816-4100
INFORMATION FOR SEQ ID NO: 42:
SEQUENCE CHARACTERISTICS:
LENGTH: 1614 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-09-060-299-42

Query Match 94.5%; Score 8259.5; DB 4; Length 1614;
Best Local Similarity 93.9%; Pred. No. 0;
Matches 1519; Conservative 41; Mismatches 51; Indels 7; Gaps 2;

Qy	1	MEAAP---PGPPWPLLLLLALLLCGCPAPAAASPLLLFANRRDRLVLDAGGVKLESTI	57
Db	1	METAPTRAPPPPPPLLLLVLSL---VFAASPLLLFANRRDRLVLDAGGVKLESTI	56
Qy	58	VVSGLEDAAAVDFQSGKAVTWDVSEBAIKQTVLNOTGAAVQNVVLSGLVSPDGLACDW	117
Db	57	VASGLEDAAAVDFQSGKAVTWDVSEBAIKQTVLNOTGAQAQIVLSGLVSPDGLACDW	116
Qy	118	VGKKLYWTDSTNRLEVANLNGTSRKVLFWDLQDQRAIALDPAHGYWYTWTDGSGTRIE	177
Db	117	VGKKLYWTDSTNRLEVANLNGTSRKVLFWDLQDQRAIALDPAHGYWYTWTDGSGAPRIE	176
Qy	178	RAGMDGSTRKLIIVDSDIYPNGLTIDLEEQKLYWADAKLSFIHRANLDGSPFRQKVVEGSL	237
Db	177	RAGMDGSTRKLIIVDSDIYPNGLTIDLEEQKLYWADAKLSFIHRANLDGSPFRQKVVEGSL	236
Qy	238	THPALTLSGDTLWTDWQTRSIHACNKTGKGKKEILSALYSMDIQVLQSERQPPFHT	297
Db	237	THPALTLSGDTLWTDWQTRSIHACNKTGKGKKEILSALYSMDIQVLQSERQPPFHT	296
Qy	298	RCEBDNGCSHLCLLSPEPYTCACPTGVQLQNGRTCKAGAEVLLARTRLRISL	357
Db	297	PCBEDNGCSHLCLLSPEPYTCACPTGVQLQNGRTCKTGAEBVLLARTRLRISL	356
Qy	358	DTDPFTDIVLQVDIRHAIAIDYDPLEGYVYTWDEVRARRAYLDGSGAQLVNTNIND	417
Db	357	DTDPFTDIVLQVDIRHAIAIDYDPLEGYVYTWDEVRARRAYLDGSGAQLVNTNIND	416
Qy	418	PDGIADWAAANLWTDGTDRIEVTRLNGTSRKILVSEDLDEPRALHPVMGLMYWTD	477
Db	417	PDGIADWAAANLWTDGTDRIEVTRLNGTSRKILVSEDLDEPRALVLPVPMGLMYWTD	476
Qy	478	WGNPKIECANLDCQERVLVNASIGWPNGLALDLQEGKLYWGDAKTDKIEVINVDGTR	537
Db	477	WGNPKIECANLDCGRHVLVNTSLGPNGLALDLQEGKLYWGDAKTDKIEVINIDGTR	536
Qy	538	RTLLEDKLPHIFGFTLLGDFIYWTWQRRSIRVHVKKASRDVIIDQLPDLMLKAVNVA	597

Search completed: February 17, 2005, 01:36:50

Db	537	KTLLLEDKLPHIFGFTLLGDFIYWTWQRRSIRVHVKKASRDVIIDQLPDLMLKAVNVA	596
Qy	598	KVGTNPCADNRGCGSHLCFFTPHATRCGCGIGLELLSDMKTCIVPEAFIPTSAAATHR	657
Db	597	KVGTNPCADNRGCGSHLCFFTPHATRCGCGIGLELLSDMKTCIPEAFIPTSAAATHR	656
Qy	658	ISLETNNNDVAIPLTGVKEASALDPDVSNHHIYWTDSLKTSISRAFNMGSSVEHVVEFGL	717
Db	657	ISLETNNNDVAIPLTGVKEASALDPDVSNHHIYWTDSLKTSISRAFNMGSSVEHVIEFGL	716
Qy	718	DYPEGMAVDMMGKNLYNADTGTNRIEVARLDQGFQVLRDLDNPRSLALDPTKGYIYW	777
Db	717	DYPEGMAVDMMGKNLYNADTGTNRIEVARLDQGFQVLRDLDNPRSLALDPTKGYIYW	776
Qy	778	TEWGGKPRIVRAFMVGTNCMTLVKVGGRANDLTIDYADORLYWTDLTNMISSNMLQGE	837
Db	777	TEWGGKPRIVRAFMVGTNCMTLVKVGGRANDLTIDYADORLYWTDLTNMISSNMLQGE	836
Qy	838	RVIADDLPHFPGLTOYSDIYWTWNLHLSIERADKTSGRNRTLIOGHLDFVMDILVPHS	897
Db	837	RVIADDLPHFPGLTOYSDIYWTWNLHLSIERADKTSGRNRTLIOGHLDFVMDILVPHS	896
Qy	898	SRQDGLNDCHMNGCCGQCLAIPEGHRCGCASHYTLDPSSRNCSPPPTFFLLFSQKSAIS	957
Db	897	SRQDGLNDCHVNSGCCGQCLAIPEGHRCGCASHYTLDPSSRNCSPPPTFFLLFSQKFAIS	956
Qy	958	RMIPDDQHSPLIILPHGLRNKAIIDYDPLDKFIYWDGRQNIKRAKDDGTQPFVLTSL	1017
Db	957	RMIPDDQLSPDLVPLHGLRNKAIINYPDLDKFIYWDGRQNIKRAKDDGTQPSMLTSP	1016
Qy	1018	QGMNPDQPHDLSIDIYSRTLFTCEATNTINVRHLSGEAMGVLRGRDRKPRAIVNAE	1077
Db	1017	QLSLSPDQPHDLSIDIYSRTLFTCEATNTINVRHLDGDMGVLRGRDRKPRAIVNAE	1076
Qy	1078	RGYLYFTNMQDRAAKIERASLDGTEREVLFTGLIRPVALVVDNLTGLKFWVDADLKRIE	1137
Db	1077	RGYLYFTNMQDRAAKIERASLDGTEREVLFTGLIRPVALVVDNLTGLKFWVDADLKRIE	1136
Qy	1138	SCDLSGANRLTLEDAIVQPLGLTILGKHLWIDROQOMIERVEKTTGDKTRIOGRAVH	1197
Db	1137	SCDLSGANRLTLEDAIVQPLGLTILGKHLWIDROQOMIERVEKTTGDKTRVQGRVTH	1196
Qy	1198	LTGTHAVEESLEBSAHPACARDNGGCSHICIAKGDGTPRCSPVHLVQLNLLTCGEP	1257
Db	1197	LTGTHAVEESLEBSAHPACARDNGGCSHICIAKGDGTPRCSPVHLVQLNLLTCGEP	1256
Qy	1258	TCSPDQFACATGEIDCIPGAWRCDFPECDQSDDEEGCPVCSAAQFPCARGQCVDLRLC	1317
Db	1257	TCSPDQFACATGEIDCIPGAWRCDFPECDQSDDEEGCPVCSAQFPCARGQCVDLRLC	1316
Qy	1318	DGEADCDRSDRDAICLBNQFRCASGQCVLIKQCCDSFDCIDGSDLMCEITKPPS	1377
Db	1317	DGEADCDRSDRDAICLBNQFRCATSGQCVLIKQCCDSFDCADGSDLMCEINKPPS	1376
Qy	1378	DDSPAHSASAIQPVIGIILSLFVMGVYVFCORVVCORVAGANGPPPHYVSGTPHVLNF	1437
Db	1377	DDIPAHSASAIQPVIGIILSLFVMGVYVFCORVVCORVYTGASGPPPHYVVGAPVPLNF	1436
Qy	1438	IAPGSGHQGPPTGIACGKSMMSVSLMGGRGVPLYDRNHVYTGASSSSSSSTKATLYPPI	1497
Db	1437	IAPGSGHQGPPTGIPCSKSVMSLSLVGGRGVPLYDRNHVYTGASSSSSSSTKATLYPPI	1496
Qy	1498	LNPPSPATDPSLXNMDMFYSSNIPATAPRPRVYIIRGMAPPTPCSDVCDSDYSASRW	1557
Db	1497	LNPPSPATDPSLXNVDVYFSSGIPATAPRPRVYIIRGMAPPTPCSDVCDSDYSISRW	1556
Qy	1558	KASKYLDLNSDSDPYPPPPPHSHQYLSAEDSCPPSPATERSYPHLPFPPSPCTDSS	1615
Db	1557	KSKYLDLNSDSDPYPPPPPHSHQYLSAEDSCPPSPCTERSYCHLFPFPPSPCTDSS	1614

Job time : 58 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 18, 2005, 20:08:43 ; Search time 2592 Seconds
(without alignments)

11545.078 Million cell updates/sec

Title: US-09-931-375A-1

Perfect score: 5063

Sequence: 1 gccatgagcccgagtgagc.....aggctgggagaaacttgta 5063

Scoring table:

IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5384158 seqs, 2955248155 residues

Total number of hits satisfying chosen parameters: 10768316

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*

2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*

3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*

4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*

5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*

6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*

7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*

8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*

9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*

10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*

11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*

12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*

13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*

14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*

15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*

16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*

17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*

18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*

19: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*

20: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*

21: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*

22: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5063	100.0	5063	10	US-09-931-375A-1
2	5058.2	99.9	5120	17	US-10-374-979-1
3	5058.2	99.9	5120	17	US-10-182-936A-1
4	5058.2	99.9	5120	18	US-10-731-739-1
5	5058.2	99.9	5120	18	US-10-477-238A-1
6	5058.2	99.9	5120	18	US-10-680-287A-1
7	5058.2	99.9	5278	18	US-10-477-238A-807
8	5058.2	99.9	5278	18	US-10-680-287A-807
9	5056.6	99.9	5120	17	US-10-374-979-2
10	5056.6	99.9	5120	17	US-10-182-936A-2
11	5056.6	99.9	5120	18	US-10-731-739-2

12	5056.6	99.9	5120	18	US-10-477-238A-2
13	5056.6	99.9	5120	18	US-10-680-287A-2
14	5056.6	99.9	5278	18	US-10-477-238A-806
15	5056.6	99.9	5278	18	US-10-680-287A-806
16	5052	99.8	5100	18	US-10-789-378-49
17	5028	99.3	5098	16	US-10-331-907-1
18	5010.8	99.0	5135	13	US-10-087-192-1409
19	4999	98.7	5059	18	US-10-723-860-3343
20	4952.6	97.8	5599	18	US-10-723-860-7329
21	4902	96.8	5022	16	US-10-331-907-32
22	4894.4	96.7	5263	16	US-10-331-907-31
23	4890.6	96.6	5125	16	US-10-331-907-26
24	4890	96.6	5166	16	US-10-331-907-23
25	4899.6	96.6	5162	16	US-10-331-907-33
26	4889	96.6	4959	16	US-10-331-907-38
27	4840	95.6	4848	17	US-10-464-368-100
28	4813	95.1	4843	16	US-10-331-907-2
29	4723.6	93.3	4915	16	US-10-331-907-28
30	4322.6	85.4	4351	16	US-10-331-907-24
31	3931	77.6	5119	9	US-09-887-540A-1
32	3912.2	77.3	5043	14	US-10-175-523-70
33	3907	77.2	5117	16	US-10-331-907-40
34	3812.6	75.3	4845	17	US-10-464-368-99
35	3788.6	74.8	4843	16	US-10-331-907-41
36	3199.6	63.2	4129	13	US-10-087-192-1406
37	2525.8	49.9	6076	17	US-10-464-368-112
38	2012.2	39.7	4842	17	US-10-464-368-38
39	1915	37.8	4842	17	US-10-464-368-39
40	1292.2	25.5	1581	17	US-10-464-368-113
41	1270.2	25.1	1711	16	US-10-331-907-35
42	1248.2	24.7	1599	16	US-10-331-907-37
43	1200.4	23.7	1226	17	US-10-182-936A-169
44	900	17.8	934	17	US-10-182-936A-170
45	781.6	15.4	4790	17	US-10-291-172-523

ALIGNMENTS

RESULT 1

US-09-931-375A-1
; Sequence 1, Application US/09931375A
; Publication No. US20030027151A1
; GENERAL INFORMATION:
; APPLICANT: WARMAN, Matthew L.
; APPLICANT: GONG, Yaqin
; APPLICANT: OLSEN, Bjorn R.
; APPLICANT: RAWADI, Georges
; APPLICANT: ROMAN-ROMAN, Sergio
; TITLE OF INVENTION: REGULATOR GENE AND SYSTEM USEFUL FOR THE DIAGNOSIS AND THERAPY OF
; TITLE OF INVENTION: OSTEOPOROSIS
; FILE REFERENCE: 38464-0004
; CURRENT APPLICATION NUMBER: US/09/931.375A
; CURRENT FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: US 60/304,851
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/234,337
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US 60/226,119
; PRIOR FILING DATE: 2000-08-18
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 5063
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-931-375A-1

Query Match 100.0%; Score 5063; DB 10; Length 5063;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5063; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GCCATGAGCCGAGTGAGCGCGCGGGCCCGTCCGCGCGGACACATGAGGCA 60

Db 4381 CAGCATGGCCCTTACAGGCATCGATCGGAAGATCCATGAGCTCCGTGAGCCTG 4440
QY 4441 ATGGGGGCGGGGGGGTGCCTCTGTACGACCGGAACACGCTCACAGGGGCTCGTCC 4500
Db 4441 ATGGGGGCGGGGGGGTGCCTCTGTACGACCGGAACACGCTCACAGGGGCTCGTCC 4500
QY 4501 AGCAGTGTCTCAGACGAGGCGACGCTGTACCGCCGATCCTGAAACCGCGCCCTCC 4560
Db 4501 AGCAGTGTCTCAGACGAGGCGACGCTGTACCGCCGATCCTGAAACCGCGCCCTCC 4560
QY 4561 CCGGCGACGAGCCCTCCCTGTACACATGACATGTTCTACTCTTCAAAATTCGGGCC 4620
Db 4561 CCGGCGACGAGCCCTCCCTGTACACATGACATGTTCTACTCTTCAAAATTCGGGCC 4620
QY 4621 ACTGCGAGACGTA CAGGCCCTACATCATTCGAGGAATGGCGCCCGACGACGCCCTGC 4680
Db 4621 ACTGCGAGACGTA CAGGCCCTACATCATTCGAGGAATGGCGCCCGACGACGCCCTGC 4680
QY 4681 AGCAGCGAGTGTGTACAGCGACTACAGCGCAGCGCGCTGGAAGCGCAGAGTACTAC 4740
Db 4681 AGCAGCGAGTGTGTACAGCGACTACAGCGCAGCGCGCTGGAAGCGCAGAGTACTAC 4740
QY 4741 CTGGATTTGAACTCGGACTCAGACCCCTATCCACCCCGACCCCGACGAGGAGTAC 4800
Db 4741 CTGGATTTGAACTCGGACTCAGACCCCTATCCACCCCGACCCCGACGAGGAGTAC 4800
QY 4801 CTGTCGGGAGGACAGTGTCCCGCCCTCGCCCGCCACCGAGAGGAGTACTTCCATCTC 4860
Db 4801 CTGTCGGGAGGACAGTGTCCCGCCCTCGCCCGCCACCGAGAGGAGTACTTCCATCTC 4860
QY 4861 TTCCCGCCCTCCCTCGTCCCGCTGACGAGTCTATCTGACCTCGCGCGGCACTCTGCG 4920
Db 4861 TTCCCGCCCTCCCTCGTCCCGCTGACGAGTCTATCTGACCTCGCGCGGCACTCTGCG 4920
QY 4921 TTCTCTGTGCCCCCTGTAATAGTTTAAATATGAAAGAAATAATATATTTATGAT 4980
Db 4921 TTCTCTGTGCCCCCTGTAATAGTTTAAATATGAAAGAAATAATATATTTATGAT 4980
QY 4981 TTAATAATAATAATAATGGATTTTAAACATGAGAAATGGAATGTGATGGGGTG 5040
Db 4981 TTAATAATAATAATAATGGATTTTAAACATGAGAAATGGAATGTGATGGGGTG 5040
QY 5041 GCGAGGCTGGGAGAACTTTGTA 5063
Db 5041 GCGAGGCTGGGAGAACTTTGTA 5063

RESULT 2

US-10-374-979-1
; Sequence 1, Application US/10374979
; Publication No. US20030219793A1
; GENERAL INFORMATION:
; APPLICANT: John P. Carulli et al.
; TITLE OF INVENTION: THE HIGH BONE MASS GENE OF 11q13.3
; FILE REFERENCE: 032796-021
; CURRENT APPLICATION NUMBER: US/10/374,979
; CURRENT FILING DATE: 2003-03-04
; PRIOR APPLICATION NUMBER: US 09/544,398
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 09/543,771
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 09/229,319
; PRIOR FILING DATE: 1999-01-13
; PRIOR APPLICATION NUMBER: US 60/071,449
; PRIOR FILING DATE: 1998-01-13
; PRIOR APPLICATION NUMBER: US 60/105,511
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 109
; SEQ ID NO 1
; LENGTH: 5120
; TYPE: DNA
; ORGANISM: Homo sapiens

US-10-374-979-1

Query Match 99.9%; Score 5058.2; DB 17; Length 5120;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 5060; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GCCATGGAAGCCCGATGAGCGCGCGGCCGCTCGCGCCCGCGGACAAATGATGAGCA 60
Db 20 GCCATGGAAGCCCGATGAGCGCGCGGCCGCTCGCGCCCGCGGACAAATGATGAGCA 79
QY 61 GCGCGCCCG 120
Db 80 GCGCGCCCG 139
QY 121 GCGTCCCG 180
Db 140 GCGTCCCG 199
QY 181 CGGCTGTGTGAGCG 240
Db 200 CGGCTGTGTGAGCG 259
QY 241 GATCG 300
Db 260 GATCG 319
QY 301 GAGGAGGCCATCAAGACGAGACTTACCTGAAACAGACGCGCGCGCGCGCGCGCG 360
Db 320 GAGGAGGCCATCAAGACGAGACTTACCTGAAACAGACGCGCGCGCGCGCGCG 379
QY 361 ATCTCCCG 420
Db 380 ATCTCCCG 439
QY 421 TGGACGCACTCAGACAGCAACCGCATCGAGTGTGCGCAACCTCAATGGCACATCCCGAAG 480
Db 440 TGGACGCACTCAGACAGCAACCGCATCGAGTGTGCGCAACCTCAATGGCACATCCCGAAG 499
QY 481 GTGCTCTTGTGGAGGACCTTGTGACGAGCTAGGGCCATTCGCTTTGGACCCCGCTCAGCGG 540
Db 500 GTGCTCTTGTGGAGGACCTTGTGACGAGCGAGGCGCATCGCTTTGGACCCCGCTCAGCGG 559
QY 541 TACATGTACTGGACAGACTGCGGTGAGACCGCGCGCGCGCGCGCGCGCGCGCGCG 600
Db 560 TACATGTACTGGACAGACTGCGGTGAGACCGCGCGCGCGCGCGCGCGCGCGCG 619
QY 601 AGCACCCGGAAGATCATTTGTGGAATCGGACATTTACTTGGCCCAATGGACATGACCATCGAC 660
Db 620 AGCACCCGGAAGATCATTTGTGGAATCGGACATTTACTTGGCCCAATGGACATGACCATCGAC 679
QY 661 CTGGAGGAGCAGAGCTCTTACTTGGGCTGACGCGCAAGCTCAGCTTTCATCCACCGTGCACAC 720
Db 680 CTGGAGGAGCAGAGCTCTTACTTGGGCTGACGCGCAAGCTCAGCTTTCATCCACCGTGCACAC 739
QY 721 CTGGACGCGCTGTTTCGCGGACAGAGTGTGTGAGGAGCGCTGACGACCGCTTCGCGCTG 780
Db 740 CTGGACGCGCTGTTTCGCGGACAGAGTGTGTGAGGAGCGCTGACGACCGCTTCGCGCTG 799
QY 781 ACGCTCTCCCGGACACTCTGTACTGGACAGACTGGCAGACCGCGCTCCATCCATGCGCTGC 840
Db 800 ACGCTCTCCCGGACACTCTGTACTGGACAGACTGGCAGACCGCGCTCCATCCATGCGCTGC 859
QY 841 AACAGCGCACTGGGGGAGAGGAGATCCTGAGTGCCTCTACTACCCCATGAC 900
Db 860 AACAGCGCACTGGGGGAGAGGAGATCCTGAGTGCCTCTACTACCCCATGAC 919
QY 901 ATCCAGGTGTGAGCCAGGAGCGCGACCTTTTCTTCCACACTCGCTGTGAGGAGGACAAT 960
Db 920 ATCCAGGTGTGAGCCAGGAGCGCGACCTTTTCTTCCACACTCGCTGTGAGGAGGACAAT 979
QY 961 GCGGCTGTCTCCCACTGTGCTGTGTGTCCTCCCAAGGAGGAGCTTTCTACATGCGCTGC 1020
Db 980 GCGGCTGTCTCCCACTGTGCTGTGTGTCCTCCCAAGGAGGAGCTTTCTACATGCGCTGC 1039

QY	1021	CCCACGGGTGTGACGCTGACGACAAACGCGAGACGCTGTAAAGGCAGGAGCCGAGGAGGTG	1080	2101	GTGTCCAAACAAACACATCTACTGGACAGACGCTGACGCTGAAAGACCATCAGCGCGCCTTC	2160
DB	1040	CCCACGGGTGTGACGCTGACGACAAACGCGAGACGCTGTAAAGGCAGGAGCCGAGGAGGTG	1099	2120	GTGTCCAAACAAACATCTACTGGACAGACGCTGACGCTGAAAGACCATCAGCGCGCCTTC	2179
QY	1081	CTGCTGTGCGCCGCGGACGACGACCTTACGAGGATCTCGCTGGACACGCCCGGACTTCACC	1140	2161	ATGAAACGGGAGCTCGGTGGAGCAAGTGTGTGAGTTTGCCCTTGACTACCCGAGGGCATG	2220
DB	1100	CTGCTGTGCGCCGCGGACGACGACCTTACGAGGATCTCGCTGGACACGCCCGGACTTCACC	1159	2180	ATGAAACGGGAGCTCGGTGGAGCAAGTGTGTGAGTTTGCCCTTGACTACCCGAGGGCATG	2239
QY	1141	GACATCGTGTGCGAGTGGACGACATCCGGACGCGCATTTGCCATCGACTACGACCCGCTA	1200	2221	GCCGTGTGACTGATGAGTGGGCAAGAACTCTACTTGGGCGGACACTTGGACCAACAGATTCGAA	2280
DB	1160	GACATCGTGTGCGAGTGGACGACATCCGGACGCGCATTTGCCATCGACTACGACCCGCTA	1219	2240	GCCGTGTGACTGATGAGTGGGCAAGAACTCTACTTGGGCGGACACTTGGACCAACAGATTCGAA	2299
QY	1201	GAGGCTATGTCTACTTGGACAGATGACGAGGTGCGGGCCATCCGAGGCGGTACTCTGGAC	1260	2281	GTGCGCGGCTGGACGCGGCACTTCGGCAAGTCTCTGTGTGGAGGACTTGGACCAACCG	2340
DB	1220	GAGGCTATGTCTACTTGGACAGATGACGAGGTGCGGGCCATCCGAGGCGGTACTCTGGAC	1279	2300	GTGCGCGGCTGGACGCGGCACTTCGGCAAGTCTCTGTGTGGAGGACTTGGACCAACCG	2359
QY	1261	GGTCTGGGGCGGACAGCCTGTGTCAACACCGAGATCAACGACCCCGATGGCATCGCGTC	1320	2341	AGTGTGTGCGCTGGATTCCTCCAAAGGGCTTACTACTTGGACCGAGTGGGGCGGCAAG	2400
DB	1280	GGTCTGGGGCGGACAGCCTGTGTCAACACCGAGATCAACGACCCCGATGGCATCGCGTC	1339	2360	AGTGTGTGCGCTGGATTCCTCCAAAGGGCTTACTACTTGGACCGAGTGGGGCGGCAAG	2419
QY	1321	GACTGGGTGGCCGGAACCTCTACTGGACCGACGAGCGGACCGGATCGAGGTGACG	1380	2401	CCGAGGATCGTGGCGGCTTCATGGACGGGACCAACTGCATGACGCTGGTGACAGGTG	2460
DB	1340	GACTGGGTGGCCGGAACCTCTACTGGACCGGACCGGATCGAGGTGACG	1399	2420	CCGAGGATCGTGGCGGCTTCATGGACGGGACCAACTGCATGACGCTGGTGACAGGTG	2479
QY	1381	CGCTCAACGCGCACCTCCCGCAAGATCTGTGTGCGAGGACCTGGACGAGCCCGAGCC	1440	2461	GGCGGGGCAACGACCTCACCATTTGACTACGCTGACAGCGCCTTACTTGGACCGACCTG	2520
DB	1400	CGCTCAACGCGCACCTCCCGCAAGATCTGTGTGCGAGGACCTGGACGAGCCCGAGCC	1459	2480	GGCGGGGCAACGACCTCACCATTTGACTACGCTGACAGCGCCTTACTTGGACCGACCTG	2539
QY	1441	ATCGCACTGCAACCCCGTGTATGGGCTCATGTACTGGACAGACTGGGGAGAGAACCTTAAA	1500	2521	GACACCAACATGATCGAGTGTCTCAACATGTGGGTGAGGACCGGCTCGTATGCCGAC	2580
DB	1460	ATCGCACTGCAACCCCGTGTATGGGCTCATGTACTGGACAGACTGGGGAGAGAACCTTAAA	1519	2540	GACACCAACATGATCGAGTGTCTCAACATGTGGGTGAGGACCGGCTCGTATGCCGAC	2599
QY	1501	ATCGAGTGTGCAACTTGGATGGGACGAGCGGGGTGTGTGTCAATGCTTCCCTCGGG	1560	2581	GATCTCCGCGACCCGTTTGGTCTGACCGGAGTACAGCGATATATCTACTTGGACAGACTGG	2640
DB	1520	ATCGAGTGTGCAACTTGGATGGGACGAGCGGGGTGTGTGTCAATGCTTCCCTCGGG	1579	2600	GATCTCCGCGACCCGTTTGGTCTGACCGGAGTACAGCGATATATCTACTTGGACAGACTGG	2659
QY	1561	TGGGCCCAACGGCCTGGACCTGCAAGGAGGAGGCTTACTTGGGGAGAGCCCAAG	1620	2641	AATCTGCACAGCATTTGACGGGCGGACAAAGACTAGCGGCGGAAACCGCACCTCATCCAG	2700
DB	1580	TGGGCCCAACGGCCTGGACCTGCAAGGAGGAGGCTTACTTGGGGAGAGCCCAAG	1639	2660	AATCTGCACAGCATTTGACGGGCGGACAAAGACTAGCGGCGGAAACCGCACCTCATCCAG	2719
QY	1621	ACAGCAAGATCGAGGTGATCAATGTTGATGGGACGAGAGCGCGGACCTCTCGAGGAC	1680	2701	GGCCACTGTGACTTTCGTGATGGACATCTCTGTGTGTTCATCTCTCCCGCAGGATGGCCTC	2760
DB	1640	ACAGCAAGATCGAGGTGATCAATGTTGATGGGACGAGAGCGCGGACCTCTCGAGGAC	1699	2720	GGCCACTGTGACTTTCGTGATGGACATCTCTGTGTGTTCATCTCTCCCGCAGGATGGCCTC	2779
QY	1681	AAGCTCCGCGCATTTTCGGGTTCACGCTGTCTGGGGACTTTCATCTACTGACTGACTGG	1740	2761	AATGACTGTATGCAACAAACGGGCGAGTGTGGGAGCTGTGCTTGGCCATCCCCGGCGGC	2820
DB	1700	AAGCTCCGCGCATTTTCGGGTTCACGCTGTCTGGGGACTTTCATCTACTGACTGACTGG	1759	2780	AATGACTGTATGCAACAAACGGGCGAGTGTGGGAGCTGTGCTTGGCCATCCCCGGCGGC	2839
QY	1741	CAGCGCCGACATCGAGCGGTGCAAGGTCAGGCGGAGCTCATCATTTGAC	1800	2821	CACCGCTGCGGCTGCGCCTCACACTACACCTTGACCCCGCAGCGCGCAACTGCAAGCCCG	2880
DB	1760	CAGCGCCGACATCGAGCGGTGCAAGGTCAGGCGGAGCTCATCATTTGAC	1819	2840	CACCGCTGCGGCTGCGCCTCACACTACACCTTGACCCCGCAGCGCGCAACTGCAAGCCCG	2899
QY	1801	CAGCTGCCGACCTGATGGGCTCAAAGCTGTGATTTGGGCAAGGTCGTGGAAACCAAC	1860	2881	CCCAACCACTTTCGTGCTTTCAGCCAGAAATCTGCCATCTAGTCGGATGATCCCCGACGAC	2940
DB	1820	CAGCTGCCGACCTGATGGGCTCAAAGCTGTGATTTGGGCAAGGTCGTGGAAACCAAC	1879	2900	CCCAACCACTTTCGTGCTTTCAGCCAGAAATCTGCCATCTAGTCGGATGATCCCCGACGAC	2959
QY	1861	CCGTGTGGGACAGGAAACGGGGGTGACGACCACTGTGCTTTCACACCCCAAGCAAC	1920	2941	CAGCAAGCCCGGATCTCATCTGCGCTGATGGACGAGTCAAGAGCCATCCGAC	3000
DB	1880	CCGTGTGGGACAGGAAACGGGGGTGACGACCACTGTGCTTTCACACCCCAAGCAAC	1939	2960	CAGCAAGCCCGGATCTCATCTGCGCTGATGGACGAGTCAAGAGCCATCCGAC	3019
QY	1921	CGGTGTGGCTGCCCATCGGCTGTGAGTGTGATGACATGAGACCTGATCGTGCCT	1980	3001	TATGACCACTTGGACAAAGTTTCATCTACTGGGTGATGGGCGCCAGAAACATCAAGCGAGCC	3060
DB	1940	CGGTGTGGCTGCCCATCGGCTGTGAGTGTGATGACATGAGACCTGATCGTGCCT	1999	3020	TATGACCACTTGGACAAAGTTTCATCTACTGGGTGATGGGCGCCAGAAACATCAAGCGAGCC	3079
QY	1981	GAGGCTTTTGGTCTTCAACAGCAGAGCCGCGCATCCACAGGATCTCCCTCGAGACCAAT	2040	3061	AAGGACGAGCGGACCCAGCCCTTTTGTGACTTCTGTAGCCCAAGGCGCAAAACCCAGAC	3120
DB	2000	GAGGCTTTTGGTCTTCAACAGCAGAGCCGCGCATCCACAGGATCTCCCTCGAGACCAAT	2059	3080	AAGGACGAGCGGACCCAGCCCTTTTGTGACTTCTGTAGCCCAAGGCGCAAAACCCAGAC	3139
QY	2041	AACAAACGAGTGGCCATCCCGCTACCGGGGTCAAGGAGGCTCAGCCCTGGACTTTGAT	2100	3121	AGGACGAGCGGACCCAGCTTGGATCGACATCTTACAGCGGACACTGTTTCTGACGTCGCGAG	3180
DB	2060	AACAAACGAGTGGCCATCCCGCTACCGGGGTCAAGGAGGCTCAGCCCTGGACTTTGAT	2119	3140	AGGACGAGCGGACCCAGCTTGGATCGACATCTTACAGCGGACACTGTTTCTGACGTCGCGAG	3199
				3181	GCCACCAATACCATCAACGCTCCACAGGCTGAGCGGGGAGCCATGGGGGTGTGTGTGCTGCT	3240

1940 CGGTGGCTGCCCATCGGCTCGAGCTGCTGAGTGACATGAACACTGATCGTGCCT 1999
1981 GAGGCTTTCTTGGTCTTACACAGCAGAGCGGCCATCCACAGGATCTCCCTCGAGACCAAT 2040
2000 GAGGCTTTCTTGGTCTTACACAGCAGAGCGGCCATCCACAGGATCTCCCTCGAGACCAAT 2059
2041 AACAAAGAGCTGGGCTATCCGCTCACCGGCGTCAAGGAGGCTCAGCCCTGGACTTTGAT 2100
2060 AACAAAGAGCTGGGCTATCCGCTCACCGGCGTCAAGGAGGCTCAGCCCTGGACTTTGAT 2119
2101 GTGTCCAAACACACATCTACTGACACAGCTCAGCTGAGACCATCAGCCGCGCTTC 2160
2120 GTGTCCAAACACACATCTACTGACACAGCTCAGCTGAGACCATCAGCCGCGCTTC 2179
2161 ATGAACGGGAGCTCGGTGGAGCAAGTGGTGAAGTTGGCTTTGACTACCCGAGGGGATG 2220
2180 ATGAACGGGAGCTCGGTGGAGCAAGTGGTGAAGTTGGCTTTGACTACCCGAGGGCATG 2239
2221 GCGTTGACTGGATGGGCAAGAACTTACTGGGCGGACACTGGACCAACAGAAATCGAA 2280
2240 GCGTTGACTGGATGGGCAAGAACTTACTGGGCGGACACTGGGACCAACAGAAATCGAA 2299
2281 GTGGCGGCTGGAGCGGCGAGTTCGGCAAGTCTCGTGTGGAGGACTTGGACCAACCCG 2340
2300 GTGGCGGCTGGAGCGGCGAGTTCGGCAAGTCTCGTGTGGAGGACTTGGACCAACCCG 2359
2341 AGGTGCTGGGCTGGATGCCCAACCAAGGGCTACATCTACTGGAGCGAGTGGGGCGGCAAG 2400
2360 AGGTGCTGGGCTGGATGCCCAACCAAGGGCTACATCTACTGGAGCGAGTGGGGCGGCAAG 2419
2401 CCGAGGATCGTGGGCTTTCATGAGCGGACCACTGATGACGCTGGTGGACAAAGTG 2460
2420 CCGAGGATCGTGGGCTTTCATGAGCGGACCACTGATGACGCTGGTGGACAAAGTG 2479
2461 GGCGGGGCAACGACTCACATTTGACTACGCTGACAGCGCTTCTACTGGACCGACCTG 2520
2480 GGCGGGGCAACGACTCACATTTGACTACGCTGACAGCGCTTCTACTGGACCGACCTG 2539
2521 GACACCAACATGATCGAGTGGTCCAAATGCTGGGTCAAGAGCGGCTGATGATGCCGAC 2580
2540 GACACCAACATGATCGAGTGGTCCAAATGCTGGGTCAAGAGCGGCTGATGATGCCGAC 2599
2581 GATCTCCGCGACCGCTTCGCTGACGACGATACGAGGATTTATCTACTGGACAGACTGG 2640
2600 GATCTCCGCGACCGCTTCGCTGACGACGATACGAGGATTTATCTACTGGACAGACTGG 2659
2641 AATCTGCACAGCATTTAGCGGCGGACCAAGACTAGCGGCGGAAACCGCAACCTCATCCAG 2700
2660 AATCTGCACAGCATTTAGCGGCGGACCAAGACTAGCGGCGGAAACCGCAACCTCATCCAG 2719
2701 GGCACCTGGATCTGATGAGACATCTGTTGTTCCACTCTCCGCGCAGGATGGCCTC 2760
2720 GGCACCTGGATCTGATGAGACATCTGTTGTTCCACTCTCCGCGCAGGATGGCCTC 2779
2761 AATGACTGATGACAAACAAAGGCGAGTGGGAGCTGTGCTTGCATCCCGCGGCGC 2820
2780 AATGACTGATGACAAACAAAGGCGAGTGGGAGCTGTGCTTGCATCCCGCGGCGC 2839
2821 CACCGCTGGGCTGGGCTTCACTACACCTTGGACCCCGAGCGGCAACTGACAGCCG 2880
2840 CACCGCTGGGCTGGGCTTCACTACACCTTGGACCCCGAGCGGCAACTGACAGCCG 2899
2881 CCCACCACTTCTGTTGTTGAGCAGAAATCTGCCATCAGTCSGATGATCCCGAGCCAG 2940
2900 CCCACCACTTCTGTTGTTGAGCAGAAATCTGCCATCAGTCSGATGATCCCGAGCCAG 2959
2941 CAGCAGCGCCGATCTCATCTCGCTGCTGACTGAGGAAAGTCAAGACCACTGCAG 3000
2960 CAGCAGCGCCGATCTCATCTCGCTGCTGACTGAGGAAAGTCAAGACCACTGCAG 3019
3001 TATGACCACTGCAGCAAGTTCTATCTGGGTGGATGGGCGCCAGAACTCAGCGAGCC 3060

3020 TATGACCACTGCAGCAAGTTCTATCTACTGGGTGGATGGGCGCCAGAACTCAAGCGAGCC 3079
3061 AAGGACGAGCGGACCCAGCCCTTTGTTGACCTCTCTGAGCCAAAGCCAAAACCCAGAC 3120
3080 AAGGACGAGCGGACCCAGCCCTTTGTTGACCTCTCTGAGCCAAAGCCAAAACCCAGAC 3139
3121 AGGACGAGCCCAACGACCTCAGCATCGACATCTACAGCGCGGACACTGTTCTCGAGCTGCGAG 3180
3140 AGGACGAGCCCAACGACCTCAGCATCGACATCTACAGCGCGGACACTGTTCTCGAGCTGCGAG 3199
3181 GCCACCAATACCATCAACGTCACAGCTGAGCGGGGAGCCATGCGGGTGGTGTCTCGT 3240
3200 GCCACCAATACCATCAACGTCACAGCTGAGCGGGGAGCCATGCGGGTGGTGTCTCGT 3259
3241 GGGGACCGGACCAAGCCAGGGCCATCGTCTCAACGCGAGGAGGTTACCTGTACTTTC 3300
3260 GGGGACCGGACCAAGCCAGGGCCATCGTCTCAACGCGAGGAGGTTACCTGTACTTTC 3319
3301 ACCAAACATGAGGACCGGGCAGCCAAAGATCGAACGCGCAGCCCTTGGACGCGACCGAGCGC 3360
3320 ACCAAACATGAGGACCGGGCAGCCAAAGATCGAACGCGCAGCCCTTGGACGCGACCGAGCGC 3379
3361 GAGTCTCTTTCACCAACCGGCTCTATCCGCCCTGTGCGCCCTGTGCTGGTGGACAAACACTG 3420
3380 GAGTCTCTTTCACCAACCGGCTCTATCCGCCCTGTGCGCCCTGTGCTGGTGGACAAACACTG 3439
3421 GGCAGGCTGTTCTGGGTGGACGCGGACCTCAAGCGCATTTGAGAGCTGTGACCTGTCAAGG 3480
3440 GGCAGGCTGTTCTGGGTGGACGCGGACCTCAAGCGCATTTGAGAGCTGTGACCTGTCAAGG 3499
3481 GCCAACCGCTGACCTTGGAGGACGCGCAACATCGTGCAGCTCTTGGGCCCTGACCATCCTT 3540
3500 GCCAACCGCTGACCTTGGAGGACGCGCAACATCGTGCAGCTCTTGGGCCCTGACCATCCTT 3559
3541 GGCAGGATCTTACTGATGACCGGACGAGCATGATCGAGCTGTGGAGGAGGAC 3600
3560 GGCAGGATCTTACTGATGACCGGACGAGCATGATCGAGCTGTGGAGGAGGAC 3619
3601 ACCGGGACCAAGCGGACTCGCATCCAGGGCGCTGCGCCCAACCTCACTGGCATCCATGCA 3660
3620 ACCGGGACCAAGCGGACTCGCATCCAGGGCGCTGCGCCCAACCTCACTGGCATCCATGCA 3679
3661 GTGAGGAAAGTCAAGCTTGGAGGATTTCTAGCCCAACCCATGTCCTGTCAGTGGTGGC 3720
3680 GTGAGGAAAGTCAAGCTTGGAGGATTTCTAGCCCAACCCATGTCCTGTCAGTGGTGGC 3739
3721 TGCTCCCAACATCTGATTTGCCAAGGATGAGGACACACGCTGCTCATGCCAGTCCAC 3780
3740 TGCTCCCAACATCTGATTTGCCAAGGATGAGGACACACGCTGCTCATGCCAGTCCAC 3799
3781 CTCGTGCTCTTCGACAAACCTGCTGACCTGTGGAGAGCGGCCACCTGCTCCCGGACCAAG 3840
3800 CTCGTGCTCTTCGACAAACCTGCTGACCTGTGGAGAGCGGCCACCTGCTCCCGGACCAAG 3859
3841 TTTGCATGTGCAACAGGGGAGATCGACTGTATTCCTCCCGGGGCTGGCGCTGTGACGGCTTT 3900
3860 TTTGCATGTGCAACAGGGGAGATCGACTGTATTCCTCCCGGGGCTGGCGCTGTGACGGCTTT 3919
3901 CCCAGTGCATGACCAAGAGCGAGGAGGCTGCGCCGCTGTGCTCCCGCCCGCAGTTC 3960
3920 CCCAGTGCATGACCAAGAGCGAGGAGGCTGCGCCGCTGTGCTCCCGCCCGCAGTTC 3979
3961 CCTGCGCGCGGGTTCAGTGTGTGGAACCTGCGCTGCGCTGCGAGCGGCGAGGACAGTGT 4020
3980 CCTGCGCGCGGGTTCAGTGTGTGGAACCTGCGCTGCGCTGCGAGCGGCGAGGACAGTGT 4039
4021 CAGGACCGCTCAGAGCGGAGCTGTGACGCGCATCTGCTGCGCCCAACAGTTCGCGTGT 4080
4040 CAGGACCGCTCAGAGCGGAGGAGTGTGACGCGCATCTGCTGCGCCCAACAGTTCGCGTGT 4099
4081 CCGAGCGCGGAGTGTGCTCTCATCAACAGCAGTGGGAGTCTCTTCCCGGACTGTATCGAC 4140
4100 CCGAGCGCGGAGTGTGCTCTCATCAACAGCAGTGGGAGTCTCTTCCCGGACTGTATCGAC 4159

1681 AAGCTCCCGCACATTTTCGGGTTTCAACGCTGCTGGGGAGCTTCACTACTGACTGACTGG 1740
1700 AAGCTCCCGCACATTTTCGGGTTTCAACGCTGCTGGGGAGCTTCACTACTGACTGACTGG 1759
1741 CAGCGCCGACGATCGAGCGGGTGACAAAGGTCAAGGCCAGCCGGGAGCTCATATTGAC 1800
1760 CAGCGCCGACGATCGAGCGGGTGACAAAGGTCAAGGCCAGCCGGGAGCTCATATTGAC 1819
1801 CAGCTGCCGACGATCGAGCGGGTGACAAAGGTCAAGGCCAGCCGGGAGCTCGTGGAAACCAAC 1860
1820 CAGCTGCCGACGATCGAGCGGGTGACAAAGGTCAAGGCCAGCCGGGAGCTCGTGGAAACCAAC 1879
1861 CCGTGTGGGACAGGAAACGGGGGGTGACGACCACTGCTGCTTCTTCAACCCCAACCAAC 1920
1880 CCGTGTGGGACAGGAAACGGGGGGTGACGACCACTGCTGCTTCTTCAACCCCAACCAAC 1939
1921 CCGTGTGGGACAGGAAACGGGGGGTGAGCTGTGAGTGAAGACTGATGAGCTGCTGCT 1980
1940 CCGTGTGGGACAGGAAACGGGGGGTGAGCTGTGAGTGAAGACTGATGAGCTGCTGCT 1999
1981 GAGGCTTCTTGGTCTTCAACAGCAGAGCGCGGCTCAAGGAGGCTCAGCCCTCGACTTTTGAT 2040
2000 GAGGCTTCTTGGTCTTCAACAGCAGAGCGCGGCTCAAGGAGGCTCAGCCCTCGAGACCAAT 2059
2041 AAACAACGAGTGGCCATCCCGCTCAACGGGCGTCAAGGAGGCTCAGCCCTCGACTTTTGAT 2100
2060 AAACAACGAGTGGCCATCCCGCTCAACGGGCGTCAAGGAGGCTCAGCCCTCGACTTTTGAT 2119
2101 GTGTCCAAACCAACATCTACTGGACAGAGCTGAGCTGGAAGCAATCAGCGCGGCTTC 2160
2120 GTGTCCAAACCAACATCTACTGGACAGAGCTGAGCTGGAAGCAATCAGCGCGGCTTC 2179
2161 ATGAACGGGAGCTCGGTGGAGCAGTGTGTGAGTTTGGCTTGTACTACCCGAGGGCATG 2220
2180 ATGAACGGGAGCTCGGTGGAGCAGTGTGTGAGTTTGGCTTGTACTACCCGAGGGCATG 2239
2221 GCCGTGTGACTGAGTGGGCAAGAACTCTACTGGGCGGACACTGGGACCAACAGAAATCGAA 2280
2240 GCCGTGTGACTGAGTGGGCAAGAACTCTACTGGGCGGACACTGGGACCAACAGAAATCGAA 2299
2281 GTGCGCGGCTGGACGGGCAAGTTCGGGCAAGTCTCTGTGTGGAGGAGCTTGGACAAACCG 2340
2300 GTGCGCGGCTGGACGGGCAAGTTCGGGCAAGTCTCTGTGTGGAGGAGCTTGGACAAACCG 2359
2341 AGTCTGCGGCTGGATCCCAACAGGCTTACATCTACTTGGACCGAGTGGGCGGCAAG 2400
2360 AGTCTGCGGCTGGATCCCAACAGGCTTACATCTACTTGGACCGAGTGGGCGGCAAG 2419
2401 CCGAGGATCGTGGCGGCTTCAATGGACGGGACCAACTGATGACGCTGGTGGACAAAGTG 2460
2420 CCGAGGATCGTGGCGGCTTCAATGGACGGGACCAACTGATGACGCTGGTGGACAAAGTG 2479
2461 GGCGGGCCAAACGACTCAACATTTGACTACGCTGACCAAGCGCTCTACTGGACCGACTG 2520
2480 GGCGGGCCAAACGACTCAACATTTGACTACGCTGACCAAGCGCTCTACTGGACCGACTG 2539
2521 GACACCAACATGATCGAGTGTGCTCAACATGCTGGGTGAGGACGGGTGCTGATGGCGAC 2580
2540 GACACCAACATGATCGAGTGTGCTCAACATGCTGGGTGAGGACGGGTGCTGATGGCGAC 2599
2581 GATCTCCCGACCGGCTTGGTCTGACGAGTACAGCGGATATATATCTACTGGACAGACTGG 2640
2600 GATCTCCCGACCGGCTTGGTCTGACGAGTACAGCGGATATATATCTACTGGACAGACTGG 2659
2641 AATCTGACAGCATTTGAGCGGGCCGACAAAGACTAGCGGCGGAAACCGACCTCATCCAG 2700
2660 AATCTGACAGCATTTGAGCGGGCCGACAAAGACTAGCGGCGGAAACCGACCTCATCCAG 2719
2701 GGCACCTTGGATTTGGATGGACATCTGTGTGTTCATCTCTCCCGCAGGATGGGCTC 2760
2720 GGCACCTTGGATTTGGATGGACATCTGTGTGTTCATCTCTCCCGCAGGATGGGCTC 2779

2761 AATGACTGTATGACAAACACGCGGAGTGTGGGAGCTGTGCCCTTGCCATCCCGCGGCGC 2820
2780 AATGACTGTATGACAAACACGCGGAGTGTGGGAGCTGTGCCCTTGCCATCCCGCGGCGC 2839
2821 CACCGCTCGGCTCGGCTCTCACTACACCTTGGACCCAGCAGCGCAACTGACGCGCG 2880
2840 CACCGCTCGGCTCGGCTCTCACTACACCTTGGACCCAGCAGCGCAACTGACGCGCG 2899
2881 CCGACCACTTCTTGTCTTTCAGCCAGAAATCTGCCATCTCGGATGATCCCGGACGAC 2940
2900 CCGACCACTTCTTGTCTTTCAGCCAGAAATCTGCCATCTCGGATGATCCCGGACGAC 2959
2941 CAGCAGCGCGGATCTCATCTGCTCCCTGATGAGCTGAGAACTCAAGACCACTCGAC 3000
2960 CAGCAGCGCGGATCTCATCTGCTCCCTGATGAGCTGAGAACTCAAGACCACTCGAC 3019
3001 TATGACCACTTTCAGCAAGTTCATCTACTGGTGTGATGGCGCCAGAACTCAAGCGAGCC 3060
3020 TATGACCACTTTCAGCAAGTTCATCTACTGGTGTGATGGCGCCAGAACTCAAGCGAGCC 3079
3061 AAGGACGAGCGGACCCAGCCCTTGTGTTGACCTCTCTGAGCCAGGCAAAACCCAGAC 3120
3080 AAGGACGAGCGGACCCAGCCCTTGTGTTGACCTCTCTGAGCCAGGCAAAACCCAGAC 3139
3121 AGGAGCGCCCAACGACTCAGCATCGACATCTACAGCCGAGCACTGTTCTGACGTCGAG 3180
3140 AGGAGCGCCCAACGACTCAGCATCGACATCTACAGCCGAGCACTGTTCTGACGTCGAG 3199
3181 GCCACCAATACCATCAACGCTCCACAGCTGAGCGGGGAGCAATGGGGGTGGTGTGCT 3240
3200 GCCACCAATACCATCAACGCTCCACAGCTGAGCGGGGAGCAATGGGGGTGGTGTGCT 3259
3241 GGGGACCGGACAGCCGAGCCGCTGCTGCTCAACCGGAGCGAGGCTGCTGACTTC 3300
3260 GGGGACCGGACAGCCGAGCCGCTGCTGCTCAACCGGAGCGAGGCTGCTGACTTC 3319
3301 ACCAACTGAGGACCGGGGAGCAAGTCAAGCGCAGCGCTTGGACCGCAACCGAGCGC 3360
3320 ACCAACTGAGGACCGGGGAGCAAGTCAAGCGCAGCGCTTGGACCGCAACCGAGCGC 3379
3361 GAGTCTCTTTCACCAACGCGCTCATCCGCTCTGTGCGCTTGGTGTGGAACAACACTG 3420
3380 GAGTCTCTTTCACCAACGCGCTCATCCGCTCTGTGCGCTTGGTGTGGAACAACACTG 3439
3421 GGGGAGCTGTTTGGTGGAGCGGACCTGAGCGCAATGAGAGCTGTGACTGTGAGG 3480
3440 GGGGAGCTGTTTGGTGGAGCGGACCTGAGCGCAATGAGAGCTGTGACTGTGAGG 3499
3481 GCCAACCGCTGACCCCTGGAGGACCGCAACATCTGTGAGCGCTCTGGGCTGACCATCTT 3540
3500 GCCAACCGCTGACCCCTGGAGGACCGCAACATCTGTGAGCGCTCTGGGCTGACCATCTT 3559
3541 GGCAGGATCTTACTGATCGACCGCAGCAGCAGATGATCGAGCGTGTGGAGAGAC 3600
3560 GGCAGGATCTTACTGATCGACCGCAGCAGCAGATGATCGAGCGTGTGGAGAGAC 3619
3601 ACCGGGACAGCGGATCTGATCCAGGCGGCTGTGCGCCCACTCACTGGGATCCATGCA 3660
3620 ACCGGGACAGCGGATCTGATCCAGGCGGCTGTGCGCCCACTCACTGGGATCCATGCA 3679
3661 GTGAGGAACTGACCTGGAGGAGTCTTACGCCCAACCATGTGCGCGCTGACAAATGGTGGC 3720
3680 GTGAGGAACTGACCTGGAGGAGTCTTACGCCCAACCATGTGCGCGCTGACAAATGGTGGC 3739
3721 TGTCTCCACATCTGTATTGCAAGGGTGTGGGACACACCGTGTCTCATGCCCCAGTCCAC 3780
3740 TGTCTCCACATCTGTATTGCAAGGGTGTGGGACACACCGTGTCTCATGCCCCAGTCCAC 3799
3781 CTGCTGTCTCGAGAACTGCTGACCTGTGGAGCGGCGCACTGCTCCCGGACGAG 3840
3800 CTGCTGTCTCGAGAACTGCTGACCTGTGGAGCGGCGCACTGCTCCCGGACGAG 3859
3841 TTTGCAATGTGCCACAGGGAGATCGACTGTATCCCGGGGCTTGGCGCTGTGTGACGGCTTT 3900

Qy	421	TGGACGAGCTCAGAGACCAACCGCATCGAGGTGGCCAACTCAATGGCACATCCCGAAG	480	Qy	1501	ATCGAGTGTGCAACTTTGGATGGGACGAGCGGGTGTGCTGGTCAATGCTCCTCCGCGG	1560
Db	440	TGGACGAGCTCAGAGACCAACCGCATCGAGGTGGCCAACTCAATGGCACATCCCGAAG	499	Db	1520	ATCGAGTGTGCAACTTTGGATGGGACGAGCGGGTGTGCTGGTCAATGCTCCTCCGCGG	1579
Qy	481	GTGCTCTTTCTGGCAGGACCTTGAACAGCTTAAGGCGCATCGCTTGGACCCCGCTCACGGG	540	Qy	1561	TGGCCCAACGGCCTTGGCCCTGGACCTTGACAGAGGGGAAGCTTACTTGGGGAGACCCCAAG	1620
Db	500	GTGCTCTTTCTGGCAGGACCTTGAACAGCTTAAGGCGCATCGCTTGGACCCCGCTCACGGG	559	Db	1580	TGGCCCAACGGCCTTGGCCCTGGACCTTGACAGAGGGGAAGCTTACTTGGGGAGACCCCAAG	1639
Qy	541	TACATGTACTGACAGACTGGGGTGAGACGCCCGGATTGAGCGGACGGGATGATGGC	600	Qy	1621	ACAGACAAGATCGAGGTGATCAATGTTGATGGGACGAAGAGCGGACCTCTCTGGAGGAC	1680
Db	560	TACATGTACTGACAGACTGGGGTGAGACGCCCGGATTGAGCGGACGGGATGATGGC	619	Db	1640	ACAGACAAGATCGAGGTGATCAATGTTGATGGGACGAAGAGCGGACCTCTCTGGAGGAC	1699
Qy	601	AGCACCCGGAAGATCATTTGTGAGACTCGGACATTTTACTTGCCCAATGGAATGACATCGAC	660	Qy	1681	AAGCTCCCGCACATTTTTCGGGTTTCAAGCTGCTGGGGGACTTTCATCTACTGAGTACTG	1740
Db	620	AGCACCCGGAAGATCATTTGTGAGACTCGGACATTTTACTTGCCCAATGGAATGACATCGAC	679	Db	1700	AAGCTCCCGCACATTTTTCGGGTTTCAAGCTGCTGGGGGACTTTCATCTACTGAGTACTG	1759
Qy	661	CTGAGGAGCAGAGCTCTACTGGCTGACGCAAGCTCAGTTTCATCCAGTGGCCAAAC	720	Qy	1741	CAGCCCGCAGCATCGAGCGGGTGCANAGTCAAGGCCAGCCGGGACGTTCATCTATTGAC	1800
Db	680	CTGAGGAGCAGAGCTCTACTGGCTGACGCAAGCTCAGTTTCATCCAGTGGCCAAAC	739	Db	1760	CAGCCCGCAGCATCGAGCGGGTGCANAGTCAAGGCCAGCCGGGACGTTCATCTATTGAC	1819
Qy	721	CTGAGCGGCTGTTCCGGCAGAGGTGTGAGGGCAGCCTGACGACCCCTTCGCCCTG	780	Qy	1801	CAGCTGCGCCGACCTGATGGGCTCAAAGCTCTGAATGTGGCAAGGTCTGTGGAAACCAAC	1860
Db	740	CTGAGCGGCTGTTCCGGCAGAGGTGTGAGGGCAGCCTGACGACCCCTTCGCCCTG	799	Db	1820	CAGCTGCGCCGACCTGATGGGCTCAAAGCTCTGAATGTGGCAAGGTCTGTGGAAACCAAC	1879
Qy	781	ACGCTCTCCGGGGACACTCTGTACTGACAGACTGGCAGACCCGCTCCATCCATGCCCTGC	840	Qy	1861	CCGTGTGCGGACAGGAAACGGGGGTGACGCCACCTGTGCTTTTTCACACCCCAACCAACC	1920
Db	800	ACGCTCTCCGGGGACACTCTGTACTGACAGACTGGCAGACCCGCTCCATCCATGCCCTGC	859	Db	1880	CCGTGTGCGGACAGGAAACGGGGGTGACGCCACCTGTGCTTTTTCACACCCCAACCAACC	1939
Qy	841	AACAAAGCGCACTGGGGGAGAGGAGAGATCTCTGAGTGCCTCTACTCAACCAATGAC	900	Qy	1921	CGGTGTGGCTGCCCATCGGCTTGAGCTGTGAGTGCATGAAGACCTGCATCGTGCT	1980
Db	860	AACAAAGCGCACTGGGGGAGAGGAGAGATCTCTGAGTGCCTCTACTCAACCAATGAC	919	Db	1940	CGGTGTGGCTGCCCATCGGCTTGAGCTGTGAGTGCATGAAGACCTGCATCGTGCT	1999
Qy	901	ATCAGGTGCTGAGCCAGGCGGAGCCTTTCTTCCACACTCGCTGTGAGGAGACAAT	960	Qy	1981	GAGGCTTTCTTGGTCTTCAACAGCAGAGCGGCCCATCCACAGATCTCCCTCGAGACCAAT	2040
Db	920	ATCAGGTGCTGAGCCAGGCGGAGCCTTTCTTCCACACTCGCTGTGAGGAGACAAT	979	Db	2000	GAGGCTTTCTTGGTCTTCAACAGCAGAGCGGCCCATCCACAGATCTCCCTCGAGACCAAT	2059
Qy	961	GCGGCTGCTCCCACTGTGCTGTGCTCCCAAGCAGAGCCTTTCTACACATGCGCCTGC	1020	Qy	2041	AACAAAGCGCTGGGCTCATCCGCTCAACGGGCGCTCAAGGAGGCTCAGCCCTGAGCTTTGAT	2100
Db	980	GCGGCTGCTCCCACTGTGCTGTGCTCCCAAGCAGAGCCTTTCTACACATGCGCCTGC	1039	Db	2060	AACAAAGCGCTGGGCTCATCCGCTCAACGGGCGCTCAAGGAGGCTCAGCCCTGAGCTTTGAT	2119
Qy	1021	CCCACGGGTGTGAGCTGACAGCAACCGCAGGAGCTGTAAAGCAGGAGCGGAGAGGTG	1080	Qy	2101	GTGTCCCAACCAACCATCTACTGACAGAGCTGAGCTGAGGAGCATCAGCGCGGCTTTC	2160
Db	1040	CCCACGGGTGTGAGCTGACAGCAACCGCAGGAGCTGTAAAGCAGGAGCGGAGAGGTG	1099	Db	2120	GTGTCCCAACCAACCATCTACTGACAGAGCTGAGCTGAGGAGCATCAGCGCGGCTTTC	2179
Qy	1081	CTGCTGTGGCCCGGCGGACGCACTACGAGGATCTCGCTGGAACGCCCGGACTTCACC	1140	Qy	2161	ATGAAACGGAGCTCGGTGGAGCAGTGTGAGTTTGGCTTTGACTACCCCGAGGCGATG	2220
Db	1100	CTGCTGTGGCCCGGCGGACGCACTACGAGGATCTCGCTGGAACGCCCGGACTTCACC	1159	Db	2180	ATGAAACGGAGCTCGGTGGAGCAGTGTGAGTTTGGCTTTGACTACCCCGAGGCGATG	2239
Qy	1141	GACATCGTGTGACAGTGGACGACATCCGGCAGCCATTGGCATCGACTACGACCCGCTA	1200	Qy	2221	GCCGTTGACTGGATGGGCAAGAACCTCTACTGGGCGGACACTGGGACCAACAGATCGAA	2280
Db	1160	GACATCGTGTGACAGTGGACGACATCCGGCAGCCATTGGCATCGACTACGACCCGCTA	1219	Db	2240	GCCGTTGACTGGATGGGCAAGAACCTCTACTGGGCGGACACTGGGACCAACAGATCGAA	2299
Qy	1201	GAGGGCTATGTCTACTGACAGATGACAGAGTGGGGCCATCCGAGGGCGTACTGAGAC	1260	Qy	2281	GTGGCGGCTGGAACGGGCAAGTTCCGGCAAGTCTCTGTGTGGAGGGAATTGGACCAACCG	2340
Db	1220	GAGGGCTATGTCTACTGACAGATGACAGAGTGGGGCCATCCGAGGGCGTACTGAGAC	1279	Db	2300	GTGGCGGCTGGAACGGGCAAGTTCCGGCAAGTCTCTGTGTGGAGGGAATTGGACCAACCG	2359
Qy	1261	GGGTCTGGGGCGAGACCTGTTCAACCGAGATCAACGACCCCGATGGCATCGCGTC	1320	Qy	2341	AGGTCTGCTGGCCCTGGATCCCAAGGGCTTACTCTGAGCCGAGTGGGGCGGCAAG	2400
Db	1280	GGGTCTGGGGCGAGACCTGTTCAACCGAGATCAACGACCCCGATGGCATCGCGTC	1339	Db	2360	AGGTCTGCTGGCCCTGGATCCCAAGGGCTTACTCTGAGCCGAGTGGGGCGGCAAG	2419
Qy	1321	GACTGGGTGGCCCGAAGACCTTACTTGGACCGACAGCGGACCGCATCGAGGTGACG	1380	Qy	2401	CCGAGGATCTGTCGGGCTTTCATGGACGGGACCAACTGTCATGACGCTGGTGACAGGTG	2460
Db	1340	GACTGGGTGGCCCGAAGACCTTACTTGGACCGACAGCGGACCGCATCGAGGTGACG	1399	Db	2420	CCGAGGATCTGTCGGGCTTTCATGGACGGGACCAACTGTCATGACGCTGGTGACAGGTG	2479
Qy	1381	CGCTCAACGGCACTCCCGCAAGATCTGTGTGAGGAGCTTGGACGAGCCCGGAGCC	1440	Qy	2461	GGCCGGGCAACGACCTTCACTTGAATACGCTGACGAGCGCTTCTACTGAGCCGACCTG	2520
Db	1400	CGCTCAACGGCACTCCCGCAAGATCTGTGTGAGGAGCTTGGACGAGCCCGGAGCC	1459	Db	2480	GGCCGGGCAACGACCTTCACTTGAATACGCTGACGAGCGCTTCTACTGAGCCGACCTG	2539
Qy	1441	ATCGCACTGCAACCCGCTGATGGGCTCATGTACTTGGACAGACTGGGAGAGAACCTTAA	1500	Qy	2521	GACACCAACATGATCGAGTCTGTCACATGTGTGGTCAAGGAGCGGTCTGATTTGCCGAC	2580
Db	1460	ATCGCACTGCAACCCGCTGATGGGCTCATGTACTTGGACAGACTGGGAGAGAACCTTAA	1519	Db	2540	GACACCAACATGATCGAGTCTGTCACATGTGTGGTCAAGGAGCGGTCTGATTTGCCGAC	2599
				Qy	2581	GATCTCCGCGACCCGCTTTCGCTGACGCACTACAGCGATTATATCTACTGGACAGACTGG	2640

Db 2600 GATCTCCGCGACCCGTTGCTGTGACGAGTATATATCTACTGGACAGACTGG 2659
QY 2641 AATCTGCAAGCATTTGAGCGGCGGACAGACTAGCGCGGAAACCGACCCCTCATCCAG 2700
Db 2660 AATCTGCAAGCATTTGAGCGGCGGACAGACTAGCGCGGAAACCGACCCCTCATCCAG 2719
QY 2701 GGCACCTGGACTTGGTATGACATCTCTGGTGTTCACATCTCTCCGCGAGGATGGCCTC 2760
Db 2720 GGCACCTGGACTTGGTATGACATCTCTGGTGTTCACATCTCTCCGCGAGGATGGCCTC 2779
QY 2761 AATGACTGATGACAAACAAAGGAGTGTGGGAGCTGTGCTTGGCCATCCCGCGCGC 2820
Db 2780 AATGACTGATGACAAACAAAGGAGTGTGGGAGCTGTGCTTGGCCATCCCGCGCGC 2839
QY 2821 CACCGCTGGGTGGCTCTACATACACCTCTGACACCCCGAGCAGCGCAACTGACAGCCG 2880
Db 2840 CACCGCTGGGTGGCTCTACATACACCTCTGACACCCCGAGCAGCGCAACTGACAGCCG 2899
QY 2881 CCCACCACTTCTGCTGTTCAGCCAGAAATCTGCCATCAGTTCGATGATCCCGAGCAG 2940
Db 2900 CCCACCACTTCTGCTGTTCAGCCAGAAATCTGCCATCAGTTCGATGATCCCGAGCAG 2959
QY 2941 CAGCAGCCCGGATCTCATCTGCTCCCTGATGACTGAGGAACGTCAGGACCATCGAC 3000
Db 2960 CAGCAGCCCGGATCTCATCTGCTCCCTGATGACTGAGGAACGTCAGGACCATCGAC 3019
QY 3001 TATGACCACTGGAACAAGTTTCACTACTGGGTGGATGGCGGCAGAAATCAAGGAGCC 3060
Db 3020 TATGACCACTGGAACAAGTTTCACTACTGGGTGGATGGCGGCAGAAATCAAGGAGCC 3079
QY 3061 AAGGACGAGGAGCCAGCCCTTTGTTTGAACCTCTCTGAGCCAGGCGCAACCCAGAC 3120
Db 3080 AAGGACGAGGAGCCAGCCCTTTGTTTGAACCTCTCTGAGCCAGGCGCAACCCAGAC 3139
QY 3121 AAGGACGCCCAAGCCTCAGCATCGACATCTACAGCCGGAACACTTCTTGACGCTGCGAG 3180
Db 3140 AAGGACGCCCAAGCCTCAGCATCGACATCTACAGCCGGAACACTTCTTGACGCTGCGAG 3199
QY 3181 GCCACCAATACATCAACGTTCAGAGCTGAGCGGGGAAGCATGGGGGTGGTGTGCTG 3240
Db 3200 GCCACCAATACATCAACGTTCAGAGGCTGAGCGGGGAAGCATGGGGGTGGTGTGCTG 3259
QY 3241 GGGGACCGGACAAAGCCAGGSCATCTGCTGCAAGCGGAGCGAGGCTACTGTACTTC 3300
Db 3260 GGGGACCGGACAAAGCCAGGSCATCTGCTGCAAGCGGAGCGAGGCTACTGTACTTC 3319
QY 3301 ACCAATGCAAGGACCGGGCAGCCAAAGATCGAACGCGCAGGCCCTGGACCGGACCGAGCGC 3360
Db 3320 ACCAATGCAAGGACCGGGCAGCCAAAGATCGAACGCGCAGGCCCTGGACCGGACCGAGCGC 3379
QY 3361 GAGGTCTCTTACACACCGGCTCATCGCCCTGTGGCCCTGTGGTGGACAAACACTG 3420
Db 3380 GAGGTCTCTTACACACCGGCTCATCGCCCTGTGGCCCTGTGGTGGACAAACACTG 3439
QY 3421 GCGAGCTGTTCTGGGTGGACGCGACCTGAAGCGCATTTGAGAGCTGTGACTGTGAGGG 3480
Db 3440 GCGAGCTGTTCTGGGTGGACGCGACCTGAAGCGCATTTGAGAGCTGTGACTGTGAGGG 3499
QY 3481 GCCAACCGCTGACCCCTGGAGGAGCCCAACATCGTGCAGCCTCTGGGCTTGACCAATCCTT 3540
Db 3500 GCCAACCGCTGACCCCTGGAGGAGCCCAACATCGTGCAGCCTCTGGGCTTGACCAATCCTT 3559
QY 3541 GCGAGCATCTTACTGATCGACCGCAGCAGCATGATCGAGCTGTGGAGAGACC 3600
Db 3560 GCGAGCATCTTACTGATCGACCGCAGCAGCATGATCGAGCTGTGGAGAGACC 3619
QY 3601 ACCGGGGAAGAGGACTCGCATCCAGGGCGGTGTGCGCCACCTCACTGGCATCCATGCA 3660
Db 3620 ACCGGGGAAGAGGACTCGCATCCAGGGCGGTGTGCGCCACCTCACTGGCATCCATGCA 3679
QY 3661 GTGGAGGAAGTCAAGCTCGAGGAGTTTCAGCCCAACCATGTGCCCTGTGAATGGTGGC 3720

Db 3680 GTGGAGGAAGTCAAGCTCGAGGAGTTCTCAGCCCAACCATGTGCCCTGTGAACAATGGTGGC 3739
QY 3721 TGCTCCCAATCTGTATTGTCGAAGGTGATGGGACACACCGGTGCTCATGCCCAGTCCAC 3780
Db 3740 TGCTCCCAATCTGTATTGTCGAAGGTGATGGGACACACCGGTGCTCATGCCAGTCCAC 3799
QY 3781 CTGCTGCTCTGCAAGAACTCTGACCTGTGTGAGAGCGCCCACTGTCTCCCGGACGAC 3840
Db 3800 CTGCTGCTCTGCAAGAACTCTGACCTGTGTGAGAGCGCCCACTGTCTCCCGGACGAC 3859
QY 3841 TTTGATGTGTCACAGGGGAGATCGACTGTATCCCGGGGCTGTGGCTGTGTGACGGCTTT 3900
Db 3860 TTTGATGTGTCACAGGGGAGATCGACTGTATCCCGGGGCTGTGGCTGTGTGACGGCTTT 3919
QY 3901 CCCAGTGCATGACACAGAGCGACGAGAGGGCTGCCCTGTGTCTCCGCGCCAGTTTC 3960
Db 3920 CCCAGTGCATGACACAGAGCGACGAGAGGGCTGCCCTGTGTCTCCGCGCCAGTTTC 3979
QY 3961 CCTCGCGCGGGGTCAAGTGTGTGACCTGCGCTGCGCTGCGAGCGGAGGAGACTGT 4020
Db 3980 CCTCGCGCGGGGTCAAGTGTGTGACCTGCGCTGCGCTGCGAGCGGAGGAGACTGT 4039
QY 4021 CAGGACCGCTCAGACGAGGGCGACTGTGACGCCATCTGCCCTGCCCAACAGTTCCGGTGT 4080
Db 4040 CAGGACCGCTCAGACGAGGTGACTGTGACGCCATCTGCTGCTGCCAACCAGTTCCGGTGT 4099
QY 4081 GCGAGCGGCTAGTGTCTCATCAAAACAGAGTGTGACTCTTCTCCCGACTGTATCGAC 4140
Db 4100 GCGAGCGGCTAGTGTCTCATCAAAACAGAGTGTGACTCTTCTCCCGACTGTATCGAC 4159
QY 4141 GCTCCCGAGAGCTCATGTGTGAATCAACAAGCGCCCTCAGACGACAGCCGCGCCAC 4200
Db 4160 GCTCCCGAGAGCTCATGTGTGAATCAACAAGCGCCCTCAGACGACAGCCGCGCCAC 4219
QY 4201 AGCAGTGCATCGGGCCCGTCAATTGGCATCATCTCTCTCTTCGTATGGGTGGTGC 4260
Db 4220 AGCAGTGCATCGGGCCCGTCAATTGGCATCATCTCTCTCTTCGTATGGGTGGTGC 4279
QY 4261 TATTTGTGTGCCAGCGGTGGTGTGCCAGGCTATGCGGGGCGCAACGGGCGCTTCCG 4320
Db 4280 TATTTGTGTGCCAGCGGTGGTGTGCCAGGCTATGCGGGGCGCAACGGGCGCTTCCG 4339
QY 4321 CACGAGTATGTACGCGGACCCCGACGCTGCCCTCTCAATTTCTAGCCCGGGGGTTC 4380
Db 4340 CACGAGTATGTACGCGGACCCCGACGCTGCCCTCTCAATTTCTAGCCCGGGGGTTC 4399
QY 4381 CAGCATGGCCCTTTCACAGGCAATCGCATGCGGAAGTCCATGATGAGTCCGTGAGCTG 4440
Db 4400 CAGCATGGCCCTTTCACAGGCAATCGCATGCGGAAGTCCATGATGAGTCCGTGAGCTG 4459
QY 4441 ATGGGGGCGGGGCGGGTCCCTGTACGACCGGAACCACTGACAGGGGCGCTTCGTC 4500
Db 4460 ATGGGGGCGGGGCGGGTCCCTGTACGACCGGAACCACTGACAGGGGCGCTTCGTC 4519
QY 4501 AGCAGCTGCTCCAGCAGGAAGCCACGCTGTACCGCGCGATCTCTGAACCGCGCGCTTC 4560
Db 4520 AGCAGCTGCTCCAGCAGGAAGCCACGCTGTACCGCGCGATCTCTGAACCGCGCGCTTC 4579
QY 4561 CCGGCCACGAGACCCCTCCCTGTGACAAATGAGCAATGTTTACTCTTCAAAATTCGCGCC 4620
Db 4580 CCGGCCACGAGACCCCTCCCTGTGACAAATGAGCAATGTTTACTCTTCAAAATTCGCGCC 4639
QY 4621 ACTCGGAGACCGTACAGGCCCTACATCTCGAGGAATGGCGCCCGACGAGCGCCCTGC 4680
Db 4640 ACTCGGAGACCGTACAGGCCCTACATCTCGAGGAATGGCGCCCGACGAGCGCCCTGC 4699
QY 4681 AGCACCGAGCTGTGTGACGAGCTACAGCGCCAGCGCTGGAAGGCGCAGAGTACTAC 4740
Db 4700 AGCACCGAGCTGTGTGACGAGCTACAGCGCCAGCGCTGGAAGGCGCAGAGTACTAC 4759
QY 4741 CTGATTTGAATCTGGAATCGAGACCCCTATCCACCCCGACCCACGCGCCACGAGCAGTAC 4800
Db 4760 CTGATTTGAATCTGGAATCGAGACCCCTATCCACCCCGACCCACGCGCCACGAGCAGTAC 4819

1381 GSCCTCAACGGCACCTCCGCAAGATCTGTGTGTCGGAGGACCTCGAGAGCCCGGAGCC 1440
1429 CECCTCAACGGCACCTCCGCAAGATCTGTGTGTCGGAGGACCTCGAGAGCCCGGAGCC 1488
1441 ATCGCACTGCACCCCGTGTATGCGGCTCATGTACTCGACAGACTGCGGAGAGAACCTTAA 1500
1489 ATCGCACTGCACCCCGTGTATGCGGCTCATGTACTCGACAGACTGCGGAGAGAACCTTAA 1548
1501 ATCGAGTGTCCAACTTGTGATGCGGAGAGAGCGGCTGTGTGCTCAATGCTTCCCTCGGG 1560
1549 ATCGAGTGTGCCAACTTGTGATGCGGAGAGAGCGGCTGTGTGCTCAATGCTTCCCTCGGG 1608
1561 TGGCCCAACGGCCTGCGCCTTGACCTGACGAGGAGGAGCTTACTTGGGGAGAGCCCAAG 1620
1609 TGGCCCAACGGCCTGCGCCTTGACCTGACGAGGAGGAGGAGCTTACTTGGGGAGAGCCCAAG 1668
1621 ACAGACAAGATCGAGGTGATCAATGTTGATGGGACGAAAGAGCGGAGCCCTCTCTGAGGAC 1680
1669 ACAGACAAGATCGAGGTGATCAATGTTGATGGGACGAAAGAGCGGAGCCCTCTCTGAGGAC 1728
1681 AAGCTCCGCGACATTTTCGGGTTACGCTGTCTGGGGGACTTCACTACTGAGACTGACTGG 1740
1729 AAGCTCCGCGACATTTTCGGGTTACGCTGTCTGGGGGACTTCACTACTGAGACTGACTGG 1788
1741 CAGCGCCGACGATCGAGCGGTGACAAAGTCAAGGTCAGCGGAGCGGAGCGTCACTCATTTGAC 1800
1789 CAGCGCCGACGATCGAGCGGTGACAAAGTCAAGGTCAGCGGAGCGGAGCGTCACTCATTTGAC 1848
1801 CAGCTGCCCGACCTGATGGGGTCAAAGCTGTGAATGTGGCCAAAGTCTGTGGAAACCAAC 1860
1849 CAGCTGCCCGACCTGATGGGGTCAAAGCTGTGAATGTGGCCAAAGTCTGTGGAAACCAAC 1908
1861 CCGTGTGGGACAGAAACGGGGGTGACGACCTGTGCTTCTTCAACCCGACGCAACC 1920
1909 CCGTGTGGGACAGAAACGGGGGTGACGACCTGTGCTTCTTCAACCCGACGCAACC 1968
1921 CCGTGTGGCTGCCCATCGGCTGAGCTGTGAGTGACATGAAGACCTGATCGTGCT 1980
1969 CCGTGTGGCTGCCCATCGGCTGAGCTGTGAGTGACATGAAGACCTGATCGTGCT 2028
1981 GAGGCTTTCTTGTCTTCAACAGCAGAGCCCGCATCCACAGGATCTCCCTCGAGACCAAT 2040
2029 GAGGCTTTCTTGTCTTCAACAGCAGAGCCCGCATCCACAGGATCTCCCTCGAGACCAAT 2088
2041 AACACGAGTGGCCATCCCGCTACGGGCGTCAAGGAGGCTCAGCCCTGAGACTTTGAT 2100
2089 AACACGAGTGGCCATCCCGCTACGGGCGTCAAGGAGGCTCAGCCCTGAGACTTTGAT 2148
2101 GTGTCCAAACCAACATCTACTGGACAGACGTCAAGCTGAAGACCATCAGCCGCGCTTC 2160
2149 GTGTCCAAACCAACATCTACTGGACAGACGTCAAGCTGAAGACCATCAGCCGCGCTTC 2208
2161 ATGAACGGGAGCTCGGTGGAGACGTTGGTGGAGTTTGGCTTGAATACCCCGAGGGCATG 2220
2209 ATGAACGGGAGCTCGGTGGAGACGTTGGTGGAGTTTGGCTTGAATACCCCGAGGGCATG 2268
2221 GCCGTTGACTGGATGGGCAAGAACTTCTACTGGGCGGACACTGGGACCAACAGAAATCGAA 2280
2269 GCCGTTGACTGGATGGGCAAGAACTTCTACTGGGCGGACACTGGGACCAACAGAAATCGAA 2328
2281 GTGGCGGGCTGGACGGGCGAGTTCCGGCAAGTCTCTGTGTGGAGGACTTGGACAAACCCG 2340
2329 GTGGCGGGCTGGACGGGCGAGTTCCGGCAAGTCTCTGTGTGGAGGACTTGGACAAACCCG 2388
2341 AGGTGCTGGCCCTGGATCCCAACGAGGGCTTACTTACTGGACCGAGTGGGGCGGCAAG 2400
2389 AGGTGCTGGCCCTGGATCCCAACGAGGGCTTACTTACTGGACCGAGTGGGGCGGCAAG 2448
2401 CCGAGGATCGGGGCTTCTATGAGACGGGACCAACTGATGACGCTGGTGACAGAGTG 2460
2449 CCGAGGATCGGGGCTTCTATGAGACGGGACCAACTGATGACGCTGGTGACAGAGTG 2508

2461 GGCGGGGCAACGACCTCACCATTTGACTACGCTGACCAAGCGCTTCTACTGGACCGACCTG 2520
2509 GGCGGGGCAACGACCTCACCATTTGACTACGCTGACCAAGCGCTTCTACTGGACCGACCTG 2568
2521 GACACCAACATGATCGAGTCTGTCCAAACATGCTGGGTCAAGGAGCGGTCGTGATTCGCGAC 2580
2569 GACACCAACATGATCGAGTCTGTCCAAACATGCTGGGTCAAGGAGCGGTCGTGATTCGCGAC 2628
2581 GATCTCCGCGACCGGTCGTGCTGACCGAGTACAGCGATTATATCTACTTGACAGACTGG 2640
2629 GATCTCCGCGACCGGTCGTGCTGACCGAGTACAGCGATTATATCTACTTGACAGACTGG 2688
2641 AATCTGCACAGCATTTGAGCGGGCGGACAAAGACTAGCGCGGAAACCGCACCTCATCCAG 2700
2689 AATCTGCACAGCATTTGAGCGGGCGGACAAAGACTAGCGCGGAAACCGCACCTCATCCAG 2748
2701 GGCCACCTTGGACTTCTGTGATGGACATCCTGTGTGTTCCACTCTCCCGCAGGATGGCCTC 2760
2749 GGCCACCTTGGACTTCTGTGATGGACATCCTGTGTGTTCCACTCTCCCGCAGGATGGCCTC 2808
2761 AATGACTGTATGACAAACAGCGGAGTGTGGGAGCTGTGCTTGCCTATCCCGGCGGC 2820
2809 AATGACTGTATGACAAACAGCGGAGTGTGGGAGCTGTGCTTGCCTATCCCGGCGGC 2868
2821 CACCGCTGCGGCTGCGCTCACACTACACCTTGACCCAGCAGCGCAACTGACAGCCG 2880
2869 CACCGCTGCGGCTGCGCTCACACTACACCTTGACCCAGCAGCGCAACTGACAGCCG 2928
2881 CCCACCACTTCTTGTCTGTTTCAGCCAGAAATCTGCCATCAGTCGGATGATCCCGAGCGAC 2940
2929 CCCACCACTTCTTGTCTGTTTCAGCCAGAAATCTGCCATCAGTCGGATGATCCCGAGCGAC 2988
2941 CAGCACAGCCCGGATCTCATCTGCTGCTGATGGATGGAGAACTGCAAGCCATCGAC 3000
2989 CAGCACAGCCCGGATCTCATCTGCTGCTGATGGATGGAGAACTGCAAGCCATCGAC 3048
3001 TATGACCCACCTGGACAAAGTTTCATCTACTGGTGGATGGGCGCCAGAAACATCAAGCGAGCC 3060
3049 TATGACCCACCTGGACAAAGTTTCATCTACTGGTGGATGGGCGCCAGAAACATCAAGCGAGCC 3108
3061 AAGAACGAGCGGACCCGAGCCCTTGTGTTGACCTCTCTGAGCCAAAGGCAAAAACCGAGAC 3120
3109 AAGAACGAGCGGACCCGAGCCCTTGTGTTGACCTCTCTGAGCCAAAGGCAAAAACCGAGAC 3168
3121 AGGACGCCCCACGACCTCAGCATCGACATCTACAGCGGACACATGTTCTGACGCTGCGAG 3180
3169 AGGACGCCCCACGACCTCAGCATCGACATCTACAGCGGACACATGTTCTGACGCTGCGAG 3228
3181 GCCACCAATACATCAACGCTCCACAGCTGAGCGGGGAAAGCATGCGGGGTGGTCTGCGT 3240
3229 GCCACCAATACATCAACGCTCCACAGCTGAGCGGGGAGCCATGCGGGGTGGTCTGCGT 3288
3241 GGGACCCGCGACAAAGCCGAGGCGCATCGTCTGTCAAACGCGGAGCGAGGGTACTCTGTACTTC 3300
3289 GGGACCCGCGACAAAGCCGAGGCGCATCGTCTGTCAAACGCGGAGCGAGGGTACTCTGTACTTC 3348
3301 ACCAACATGAGGAGCGGGGAGCCAAAGATCGAAACGCGGAGCCCTGGAACCGGAGCGC 3360
3349 ACCAACATGAGGAGCGGGGAGCCAAAGATCGAAACGCGGAGCCCTGGAACCGGAGCGC 3408
3361 GAGGTCTCTTACACACCGGCTCATCGCCCTGTGCGCTGTGGTGGTGGACAAACACTG 3420
3409 GAGGTCTCTTACACACCGGCTCATCGCCCTGTGCGCTGTGGTGGTGGTGGTGGTGGTGGTGG 3468
3421 GGCAGGCTGTTCTGGGTGGACCGGGAACCTGAAAGCGCATTTGAGAGCTGTGACCTGTGAGGG 3480
3469 GGCAGGCTGTTCTGGGTGGACCGGGAACCTGAAAGCGCATTTGAGAGCTGTGACCTGTGAGGG 3528
3481 GCCAACCGCTGACCTTGGAGGAGCGCAACATCTGTGAGGCTCTTGGGCTGACCATCTT 3540
3529 GCCAACCGCTGACCTTGGAGGAGCGCAACATCTGTGAGGCTCTTGGGCTGACCATCTT 3588
3541 GGCAAGCATCTCTACTGGATCGACCGGCGAGCAGCAGATGATCGAGCGTGTGGAGAGACC 3600

QY 121 GGTCTCCCGGCCCCCGCGCGCTCGCGCTCTGCTATTTGCCAACCGCGGACGTA 180
Db 129 GGTCTCCCGGCCCCCGCGCGCTCGCGCTCTGCTATTTGCCAACCGCGGACGTA 228
QY 181 CGGCTGTGTGGACCGCGCGGAGTCAAGCTGAGTCCACCATCGTGGTCAAGCGGCTGGAG 240
Db 229 CGGCTGTGTGGACCGCGCGGAGTCAAGCTGAGTCCACCATCGTGGTCAAGCGGCTGGAG 288
QY 241 GATGGCGCGGAGTGGACCTTCCAGTTTCCAAAGGAGCGGTGTACTGACAGAGTGGAGC 300
Db 289 GATGGCGCGGAGTGGACCTTCCAGTTTCCAAAGGAGCGGTGTACTGACAGAGTGGAGC 348
QY 301 GAGGAGGCGCATCAAGCAGACCTTACCTGAACACAGAGCGGCGCGCTGACAGACGTGTGC 360
Db 349 GAGGAGGCGCATCAAGCAGACCTTACCTGAACACAGAGCGGCGCGCTGACAGAGCGTGTGC 408
QY 361 ATCTCGGCGCTGTCTCTCCGACCGGCTCGCTCGGACTGGGTGGGCAAGAGCTGTAC 420
Db 409 ATCTCGGCGCTGTCTCTCCGACCGGCTCGCTCGGACTGGGTGGGCAAGAGCTGTAC 468
QY 421 TGGACGGACTCAGAGACCAACCGCATCGAGTGGCCAACTCAATGGCACAATCCCGGAAG 480
Db 469 TGGACGGACTCAGAGACCAACCGCATCGAGTGGCCAACTCAATGGCACAATCCCGGAAG 528
QY 481 GTGCTCTTCTGGCAGGACCTTGAACAGCCTAGGGCCATCGGCTTGGACCCCGCTCACGGG 540
Db 529 GTGCTCTTCTGGCAGGACCTTGAACAGCCTAGGGCCATCGGCTTGGACCCCGCTCACGGG 588
QY 541 TACATGTACTGGACAGACTGGGTGAGACGCGCCGGATTGAGCGGGCAGGGATGGATGGC 600
Db 589 TACATGTACTGGACAGACTGGGTGAGACGCGCCGGATTGAGCGGGCAGGGATGGATGGC 648
QY 601 AGCACCGGAAGATCATTTGTGGACTCGGACATTTTACTGGGCCAATGGAATGACCATGGAC 660
Db 649 AGCACCGGAAGATCATTTGTGGACTCGGACATTTTACTGGGCCAATGGAATGACCATGGAC 708
QY 661 CTGAGAGGACAGAGCTCTACTGGGTGACGCCCGGATTGAGCGGGCAGGGATGGATGGC 720
Db 709 CTGAGAGGACAGAGCTCTACTGGGTGACGCCCGGATTGAGCGGGCAGGGATGGATGGC 768
QY 721 CTGAGCGGCTCTTCCGCGAGAAAGTGGTGGAGGCGAGCTTGACGCAACCTTCGCCCTG 780
Db 769 CTGAGCGGCTCTTCCGCGAGAAAGTGGTGGAGGCGAGCTTGACGCAACCTTCGCCCTG 828
QY 781 ACGCTCTCCGGGACACTCTGTAATGGACAGACTGGGAGACCGCTCCATCCATGCTGCG 840
Db 829 ACGCTCTCCGGGACACTCTGTAATGGACAGACTGGGAGACCGCTCCATCCATGCTGCG 888
QY 841 AACAGCGCACTGGGGGGAAGAGAGATCTGTAGTGCCCTCTACTCAACCCATGGAC 900
Db 889 AACAGCGCACTGGGGGGAAGAGAGATCTGTAGTGCCCTCTACTCAACCCATGGAC 948
QY 901 ATCCAGGTGTGACGAGGAGCGGAGCTTTTCCACACTCGCTGTGAGGAGGACAAT 960
Db 949 ATCCAGGTGTGACGAGGAGCGGAGCTTTTCCACACTCGCTGTGAGGAGGACAAT 1008
QY 961 GGGCGTGTCTCCACCTGCTGCTGCTCCCAAGCGAGCTTTCTACATGCGCTGCG 1020
Db 1009 GGGCGTGTCTCCACCTGCTGCTGCTCCCAAGCGAGCTTTCTACATGCGCTGCG 1068
QY 1021 CCCACGGGTGTGACGCTGACGAGCAACGGGAGGACGTGTAGGAGGAGCGGAGGAGGTG 1080
Db 1069 CCCACGGGTGTGACGCTGACGAGCAACGGGAGGAGCTGTAGGAGGAGCGGAGGAGGTG 1128
QY 1081 CTGCTGTGTCGCGCGGAGAGGACCTTACGAGGAGATCTGCTGTGACAGCGCGGACTTCACC 1140
Db 1129 CTGCTGTGTCGCGCGGAGAGGACCTTACGAGGAGATCTGCTGTGACAGCGCGGACTTCACC 1188
QY 1141 GACATCGTGTGACAGTGGAGGACATCCGGCAGCCCAATTCGACTGAGACCGCGTA 1200
Db 1189 GACATCGTGTGACAGTGGAGGACATCCGGCAGCCCAATTCGACTGAGACCGCGTA 1248

QY 1201 GAGGCTATGTCTACTGGACAGATGACGAGGTGCGGGCCATCCGACAGGCGGTACTCTGAC 1260
Db 1249 GAGGCTATGTCTACTGGACAGATGACGAGGTGCGGGCCATCCGACAGGCGGTACTCTGAC 1308
QY 1261 GGGTCTGCGGGCGCAGACGCTGGTCAAACACGAGATCAACGACCCCGATGGCATCGCGTTC 1320
Db 1309 GGGTCTGCGGGCGCAGACGCTGGTCAAACACGAGATCAACGACCCCGATGGCATCGCGTTC 1368
QY 1321 GACTGGGTGGCCCCGAAACCTCTACTGGACCAACACGCGCACGACCCGATCGAGGTGACG 1380
Db 1369 GACTGGGTGGCCCCGAAACCTCTACTGGACCAACACGCGCACGACCCGATCGAGGTGACG 1428
QY 1381 CGGCTCAACGCGACCTCCCGCAAGATCTCTGTGTGCGAGGACCTTGGACGAGCCCGGAGCC 1440
Db 1429 CGGCTCAACGCGACCTCCCGCAAGATCTCTGTGTGCGAGGACCTTGGACGAGCCCGGAGCC 1488
QY 1441 ATCCGACTGCAACCCCGTGTATGGGCTCATGTACTGGACAGACTGGGAGAGAACCTTAAA 1500
Db 1489 ATCCGACTGCAACCCCGTGTATGGGCTCATGTACTGGACAGACTGGGAGAGAACCTTAAA 1548
QY 1501 ATCGAGTGTGCAACTTTGGATGGGACGCGGTGTGTGCTCAATGCTCTCCTCGGG 1560
Db 1549 ATCGAGTGTGCAACTTTGGATGGGACGCGGTGTGTGCTCAATGCTCTCCTCGGG 1608
QY 1561 TGGCCCCAACGSCCTGGCCCTGGACCTTCGAGAGGAGGAGCTTACTGGGAGAGCGCAAG 1620
Db 1609 TGGCCCCAACGSCCTGGCCCTGGACCTTCGAGAGGAGGAGCTTACTGGGAGAGCGCAAG 1668
QY 1621 ACAGCAAGATTCGAGGTGATCAATGTTATGGGACGAGAGCGGACCTCTCTCGAGGAC 1680
Db 1669 ACAGCAAGATTCGAGGTGATCAATGTTATGGGACGAGAGCGGACCTCTCTCGAGGAC 1728
QY 1681 AAGCTCCGCACTTTTCGGGTTCAACGCTGTGGGGGACTTCTACTGTGACTGACTGG 1740
Db 1729 AAGCTCCGCACTTTTCGGGTTCAACGCTGTGGGGGACTTCTACTGTGACTGACTGG 1788
QY 1741 CAGCGCGCAGCATCGAGCGGCTGCAAGGTCAAGGCGAGCGGAGCTTCACTATTGAC 1800
Db 1789 CAGCGCGCAGCATCGAGCGGCTGCAAGGTCAAGGCGAGCGGAGCTTCACTATTGAC 1848
QY 1801 CAGCTGCCGCACTGTATGGGCTCAAAGCTGTGTAATGTGGCAAGGTCTGTCGAGCAAC 1860
Db 1849 CAGCTGCCGCACTGTATGGGCTCAAAGCTGTGTAATGTGGCAAGGTCTGTCGAGCAAC 1908
QY 1861 CCGTGTGGGACAGGAAACGGGGGTGACGCCACTGTGCTTCTTCAACCCGACGCAAC 1920
Db 1909 CCGTGTGGGACAGGAAACGGGGGTGACGCCACTGTGCTTCTTCAACCCGACGCAAC 1968
QY 1921 CCGTGTGGCTGCCCCATCGGCTTGGAGCTGTGAGTGCATGAAGACTGTGATCGTCT 1980
Db 1969 CCGTGTGGCTGCCCCATCGGCTTGGAGCTGTGAGTGCATGAAGACTGTGATCGTCT 2028
QY 1981 GAGGCTTCTTGTGTTTCAACGAGAGCGGCCATCCAGAGGATCTCCTCGAGACCAAT 2040
Db 2029 GAGGCTTCTTGTGTTTCAACGAGAGCGGCCATCCAGAGGATCTCCTCGAGACCAAT 2088
QY 2041 AACAAAGAGCTGGGCACTCCGCTCAAGGGGTCAAGAGGCGCTCAGCCCTGGACTTTGAT 2100
Db 2089 AACAAAGAGCTGGGCACTCCGCTCAAGGGGTCAAGAGGCGCTCAGCCCTGGACTTTGAT 2148
QY 2101 GTGTCCAAACACCATCTACTGACAGAGCTCAGCTGAAGACCATCAGCCGCGCTTC 2160
Db 2149 GTGTCCAAACACCATCTACTGACAGAGCTCAGCTGAAGACCATCAGCCGCGCTTC 2208
QY 2161 ATGAAACGGGAGCTCGGTGGAGCACTGTGGGCTTTGGCTTTGACTACCCCGAGGCGATG 2220
Db 2209 ATGAAACGGGAGCTCGGTGGAGCACTGTGGGCTTTGGCTTTGACTACCCCGAGGCGATG 2268
QY 2221 GCGTGTGACTGGATGGCAAGAACCTTCACTGGGCGGACACTGGGACCAACAGATCGAA 2280
Db 2269 GCGTGTGACTGGATGGCAAGAACCTTCACTGGGCGGACACTGGGACCAACAGATCGAA 2328
QY 2281 GTGGCGGCTGGACCGGCAAGTTTCCGCAAGTCTCTGTGTGGGGGACTTTGGACAAACCG 2340

Db 1100 CTGCTGTCGGCCGCGACGGACCTACGAGGATCTCGTGGACACGCGGACTTCACC 1159
QY 1141 GACATCGTGTGAGGTGGAGCAGATCCGGCAGCCCATTTGCCATCGACTAGACCCGCTA 1200
Db 1160 GACATCGTGTGAGGTGGAGCAGATCCGGCAGCCCATTTGCCATCGACTAGACCCGCTA 1219
QY 1201 GAGGGCTATGTCTACTGGACAGATGACGAGGTGGGGCCATCCGACAGGGCGGTACTCGAC 1260
Db 1220 GAGGGCTATGTCTACTGGACAGATGACGAGGTGGGGCCATCCGACAGGGCGGTACTCGAC 1279
QY 1261 GGGTCTGGGGCGCAGACGCTGGTCAACACCGAGATCAACGACCCCGATGGCATCGCGGTC 1320
Db 1280 GGGTCTGGGGCGCAGACGCTGGTCAACACCGAGATCAACGACCCCGATGGCATCGCGGTC 1339
QY 1321 GACTGGGTGGCCGAAACCTTACTTGGACCGACACGCGGACCGGATCGAGGTGACG 1380
Db 1340 GACTGGGTGGCCGAAACCTTACTTGGACCGACACGCGGACCGGATCGAGGTGACG 1399
QY 1381 CGCCTCAACGGCACCTCCCGCAAGATCCTGTGTGCGAGGACCTGGACGAGCCCGGAGCC 1440
Db 1400 CGCCTCAACGGCACCTCCCGCAAGATCCTGTGTGCGAGGACCTGGACGAGCCCGGAGCC 1459
QY 1441 ATCGCACTGCAACCCGCTGATGGGCTCATGTACTGGACAGACTGGGAGAGAACCTTAAA 1500
Db 1460 ATCGCACTGCAACCCGCTGATGGGCTCATGTACTGGACAGACTGGGAGAGAACCTTAAA 1519
QY 1501 ATCGAGTGTGCCAATTTGGATGGGACGAGGCGGCTGTGTGTCATGCTCCCTCGGG 1560
Db 1520 ATCGAGTGTGCCAATTTGGATGGGACGAGGCGGCTGTGTGTCATGCTCCCTCGGG 1579
QY 1561 TGGCCCAACGGCCTGGCCCTGGACCTGCGAGAGGGAGGCTTACTGGGAGAGCCCAAG 1620
Db 1580 TGGCCCAACGGCCTGGCCCTGGACCTGCGAGAGGGAGGCTTACTGGGAGAGCCCAAG 1639
QY 1621 ACAGACAAGATCGAGGTGATCAATGTTGATGGGACGAGAGGCGGACCCCTCTGGAGGAC 1680
Db 1640 ACAGACAAGATCGAGGTGATCAATGTTGATGGGACGAGAGGCGGACCCCTCTGGAGGAC 1699
QY 1681 AAGCTCCGCGACATTTTCGGGTTTACGCTGTCTGGGGAGCTTCACTACTGGACGTACTGG 1740
Db 1700 AAGCTCCGCGACATTTTCGGGTTTACGCTGTCTGGGGAGCTTCACTACTGGACGTACTGG 1759
QY 1741 CAGCGCCGACATCGAGCGGGTGCACAAGGTCAAGGCGGAGGAGGCTCATCATTTGAC 1800
Db 1760 CAGCGCCGACATCGAGCGGGTGCACAAGGTCAAGGCGGAGGAGGCTCATCATTTGAC 1819
QY 1801 CAGCTGCCCGACCTGATGGGCTCAAAGCTGTGAATGTGGCCAAAGGTGCTCGGAACCAAC 1860
Db 1820 CAGCTGCCCGACCTGATGGGCTCAAAGCTGTGAATGTGGCCAAAGGTGCTCGGAACCAAC 1879
QY 1861 CCGTGTGGGACAGAAACGGGGGTGCAGCCACTGTGTCTTTTCAACCCCGACGAAAC 1920
Db 1880 CCGTGTGGGACAGAAACGGGGGTGCAGCCACTGTGTCTTTTCAACCCCGACGAAAC 1939
QY 1921 CGGTGTGGCTGCCCATCGGCTCGAGCTGTGAGTGCATGAAGACCTGCTGCTGCT 1980
Db 1940 CGGTGTGGCTGCCCATCGGCTCGAGCTGTGAGTGCATGAAGACCTGCTGCTGCT 1999
QY 1981 GAGGCTTTCTTGGTCTTCAACAGCAGAGCCCGCCATCCACAGGATCTCCCTCGAGACCAAT 2040
Db 2000 GAGGCTTTCTTGGTCTTCAACAGCAGAGCCCGCCATCCACAGGATCTCCCTCGAGACCAAT 2059
QY 2041 AACAAAGAGTGGGCTATCCGCTCAACGGGCTCAAGGAGGCTCAGGCTGGATCTTGGAT 2100
Db 2060 AACAAAGAGTGGGCTATCCGCTCAACGGGCTCAAGGAGGCTCAGGCTGGATCTTGGAT 2119
QY 2101 GTGTCCCAACACACATCTACTGACAGAGCTCAGCTGAGGACCATCAGCGCGCTTC 2160
Db 2120 GTGTCCCAACACACATCTACTGACAGAGCTCAGCTGAGGACCATCAGCGCGCTTC 2179
QY 2161 ATGAACGGGAGCTCGGTGGAGCAGCTGTGTGAGTTTGGGCTTGTACTACCCCGAGGGCATG 2220

Db 2180 ATGAACGGGAGCTCGGTGGAGCAGCTGGTGGAGTTTGGCCCTTGACTACCCCGAGGGCATG 2239
QY 2221 GCCGTTCACCTGGATGGGCAAGAACCTTACTGGGCCGACACTGGGACCAACAGAAATCGAA 2280
Db 2240 GCCGTTCACCTGGATGGGCAAGAACCTTACTGGGCCGACACTGGGACCAACAGAAATCGAA 2299
QY 2281 GTGCGCGGCTGGACGGGCGGATTCGCGCAAGTCTCGTGTGGAGGGAGCTTGGACAAACCCG 2340
Db 2300 GTGCGCGGCTGGACGGGCGGATTCGCGCAAGTCTCGTGTGGAGGGAGCTTGGACAAACCCG 2359
QY 2341 AGTTCGCTGGCCCTGGATTCGCAACAGGGCTCATCTACTGGACCGAGTGGGGCGGCAAG 2400
Db 2360 AGTTCGCTGGCCCTGGATTCGCAACAGGGCTCATCTACTGGACCGAGTGGGGCGGCAAG 2419
QY 2401 CCGAGGATCTGGCGGCTTCATGGACGGGACCAACTGCGATGACGCTGGTGGACAGGTG 2460
Db 2420 CCGAGGATCTGGCGGCTTCATGGACGGGACCAACTGCGATGACGCTGGTGGACAGGTG 2479
QY 2461 GCGCGGCGCAACGACCTCAGCATTTGACTACGCTGACGAGCGCTTCTACTGGACCGACCTG 2520
Db 2480 GCGCGGCGCAACGACCTCAGCATTTGACTACGCTGACGAGCGCTTCTACTGGACCGACCTG 2539
QY 2521 GACACCAACATGATCGAGTCTGTCACATGCTGGGTCAAGAGCGGGTCTGTGATTCGCCAC 2580
Db 2540 GACACCAACATGATCGAGTCTGTCACATGCTGGGTCAAGAGCGGGTCTGTGATTCGCCAC 2599
QY 2581 GATCTCCGCGACCCGTTCCGTTCTGACGAGTACAGCGATATATATCTACTGGACAGACTGG 2640
Db 2600 GATCTCCGCGACCCGTTCCGTTCTGACGAGTACAGCGATATATATCTACTGGACAGACTGG 2659
QY 2641 AATCTGCAACGACCTTGAAGCGGCGGCAACAGACTAGCGGCGGGAACCGCACCTCATCCAG 2700
Db 2660 AATCTGCAACGACCTTGAAGCGGCGGCAACAGACTAGCGGCGGGAACCGCACCTCATCCAG 2719
QY 2701 GGGCACCTGACCTTCTGTGATGGACATCTCGTGTTCATCTCTCCGCGCAGGATGGGCTC 2760
Db 2720 GGGCACCTGACCTTCTGTGATGGACATCTCGTGTTCATCTCTCCGCGCAGGATGGGCTC 2779
QY 2761 AATGACTGTATGCACAAACGCGGAGTGGGAGCTGTGCGCTTGCATCCCGCGGCGG 2820
Db 2780 AATGACTGTATGCACAAACGCGGAGTGGGAGCTGTGCGCTTGCATCCCGCGGCGG 2839
QY 2821 CACCGCTGCGGCTGCGCTCACACTACACCTTGACCCCGACGAGCGGCAACTGACGCGCG 2880
Db 2840 CACCGCTGCGGCTGCGCTCACACTACACCTTGACCCCGACGAGCGGCAACTGACGCGCG 2899
QY 2881 CCCACCACTTCTGTGTTTCAAGCAAAATCTGCGCATCTAGTCCGATGATCCCGGACGAC 2940
Db 2900 CCCACCACTTCTGTGTTTCAAGCAAAATCTGCGCATCTAGTCCGATGATCCCGGACGAC 2959
QY 2941 CAGCAGCGCGGATCTCATCTGCGCTGCGCTGAGTGGAGACGTCAGGACGTCAGGACGATCGAC 3000
Db 2960 CAGCAGCGCGGATCTCATCTGCGCTGCGCTGAGTGGAGACGTCAGGACGTCAGGACGATCGAC 3019
QY 3001 TATGACCCACTTGGACAAAGTTTCATCTACTGGTGGATGGGGCGGCAAGCAATCAAGCGAGCC 3060
Db 3020 TATGACCCACTTGGACAAAGTTTCATCTACTGGTGGATGGGGCGGCAAGCAATCAAGCGAGCC 3079
QY 3061 AAGSACGACGGGACCCAGCCCTTTGTTTGAACCTCTCTGAGCCCAAGGCAAAACCCAGAC 3120
Db 3080 AAGSACGACGGGACCCAGCCCTTTGTTTGAACCTCTCTGAGCCCAAGGCAAAACCCAGAC 3139
QY 3121 AGGAGCGCCACGACCTCAGCATTCGATGACCGGACGACCTGTTCTGGACGTCGAG 3180
Db 3140 AGGAGCGCCACGACCTCAGCATTCGATGACCGGACGACCTGTTCTGGACGTCGAG 3199
QY 3181 GCCACCAATACCACTCAACGCTCAGCGGGGAAGCCATGGGGGTGGTCTGGT 3240
Db 3200 GCCACCAATACCACTCAACGCTCAGCGGGGAAGCCATGGGGGTGGTCTGGT 3259
QY 3241 GGGSACCGGACCAAGCCCGAGGGCCATCGTCTGTCACCGGAGCGAGGGTACCTGTACTTC 3300
Db 3260 GGGSACCGGACCAAGCCCGAGGGCCATCGTCTGTCACCGGAGCGAGGGTACCTGTACTTC 3319

3001 TATGACCCACATGGACAAGTTCACTTACTGGGTGGATGGCGCCAGAAATCAAGCGAGCC 3060
3020 TATGACCCACATGGACAAGTTCACTTACTGGGTGGATGGCGCCAGAAATCAAGCGAGCC 3079
3061 AAGGACGACGGAGCCAGCCCTTTGTTGACCTCTCTGAGCGAAGGCGCAAAACCCAGAC 3120
3080 AAGGACGACGGAGCCAGCCCTTTGTTGACCTCTCTGAGCGAAGGCGCAAAACCCAGAC 3139
3121 AGGACGACCCACGACCTCAGATCGACATCTACAGCCGGAACATGTTCTTGAAGTGCAG 3180
3140 AGGACGACCCACGACCTCAGATCGACATCTACAGCCGGAACATGTTCTTGAAGTGCAG 3199
3181 GCCACCAATACATCAACGCTCCACAGGCTGAGCGGGGAAGCATGGGGGTGGTCTGCGT 3240
3200 GCCACCAATACATCAACGCTCCACAGGCTGAGCGGGGAAGCATGGGGGTGGTCTGCGT 3259
3241 GGGGACCGCGACAGCCCGAGGCGCATCTGTCGTCACCGGAGCGAGGTAATCTGTAATTC 3300
3260 GGGGACCGCGACAGCCCGAGGCGCATCTGTCGTCACCGGAGCGAGGTAATCTGTAATTC 3319
3301 ACCAACATGACGAGCCGGGAGCCAGATCGAAGATCGAAGCGGAGCCCTGGACCGGAGCGC 3360
3320 ACCAACATGACGAGCCGGGAGCCAGATCGAAGCGGAGCCCTGGACCGGAGCGC 3379
3361 CAGGTCTCTTCAACACCGGCTCATCCGCTGTCGCTGGGCTGGTGGACAAACACATG 3420
3380 GAGGTCTCTTCAACACCGGCTCATCCGCTGTCGCTGGGCTGGTGGACAAACACATG 3439
3421 GGCAGCTGTTCTGGGTGGACCGCGACTGAGGCGCATTTGAGAGCTGTGACCTGTGAGG 3480
3440 GGCAGCTGTTCTGGGTGGACCGCGACTGAGGCGCATTTGAGAGCTGTGACCTGTGAGG 3499
3481 GCCAACCGCTGACCCCTGGAGGACCGCAACATCGTGGAGCCCTCTGGGCTGACCATCTT 3540
3500 GCCAACCGCTGACCCCTGGAGGACCGCAACATCGTGGAGCCCTCTGGGCTGACCATCTT 3559
3541 GGCAGCATCTTACTGATCGACCGCCAGCAGCAGATGATCGAGCGTGTGGAGAGACC 3600
3560 GGCAGCATCTTACTGATCGACCGCCAGCAGCAGATGATCGAGCGTGTGGAGAGACC 3619
3601 ACCGGGACAAAGCGACTCGCATCCAGGCGCTGTGCGCCACCTCACTGGCATCCATGCA 3660
3620 ACCGGGACAAAGCGACTCGCATCCAGGCGCTGTGCGCCACCTCACTGGCATCCATGCA 3679
3661 GTGAGGAGTCAAGCTGGAGGTTCTCAGCCACCATGTGCGCGTGGACAAATGGTGGC 3720
3680 GTGAGGAGTCAAGCTGGAGGTTCTCAGCCACCATGTGCGCGTGGACAAATGGTGGC 3739
3721 TGCTCCACATCTGATTTGCAAGGTTGATGGGACCAACCATGCTCATGCGCCAGTCCAC 3780
3740 TGCTCCACATCTGATTTGCAAGGTTGATGGGACCAACCATGCTCATGCGCCAGTCCAC 3799
3781 CTGCTGCTCTGCAAACTGCTGACCTGTGAGAGCGCGCCACCTGCTCCCGGAGCCAG 3840
3800 CTGCTGCTCTGCAAACTGCTGACCTGTGAGAGCGCGCCACCTGCTCCCGGAGCCAG 3859
3841 TTTGATGTCGCAAGGGAGATCGACTGATATCCCGGGGCTGGCGCTGTGACGGCTTT 3900
3860 TTTGATGTCGCAAGGGAGATCGACTGATATCCCGGGGCTGGCGCTGTGACGGCTTT 3919
3901 CCCGAGTGCATGATACAGAGCGACGAGGAGGCTGCGCCGCTGTGCTCCCGCCCGAGTTT 3960
3920 CCCGAGTGCATGATACAGAGCGACGAGGAGGCTGCGCCGCTGTGCTCCCGCCCGAGTTT 3979
3961 CCCTGCGCGGGGTCAAGTGTGTGACCTGCGCTGCGAGCGGAGGAGTGTG 4020
3980 CCCTGCGCGGGGTCAAGTGTGTGACCTGCGCTGCGAGCGGAGGAGTGTG 4039
4021 CAGGACCGCTCAGACGAGGCGGATGTGACCGCATCTGCTGCGCCACAGTTCGGTGT 4080
4040 CAGGACCGCTCAGACGAGGCGGATGTGACCGCATCTGCTGCGCCACAGTTCGGTGT 4099

RESULT 12

US-10-477-238A-2

; Sequence 2, Application US/10477238A

; Publication No. US20040221326A1

4081 GCGAGCGCCAGTGTGCTCTCATCAAAACAGCAGTGCAGTCTCTTCCCGACATGTATCGAC 4140
4100 GCGAGCGCCAGTGTGCTCTCATCAAAACAGCAGTGCAGTCTCTTCCCGACATGTATCGAC 4159
4141 GGTCCCGACGAGCTCATGTGTGAATAACCAAGCGCCCTCAGACGACAGCGCCGCCAC 4200
4160 GGTCCCGACGAGCTCATGTGTGAATAACCAAGCGCCCTCAGACGACAGCGCCGCCAC 4219
4201 AGCAGTGCATCGGCGCCGCTCATTTGGCATCATCTCTCTCTCTCTCTCTCTCTCTCTCT 4260
4220 AGCAGTGCATCGGCGCCGCTCATTTGGCATCATCTCTCTCTCTCTCTCTCTCTCTCTCT 4279
4261 TATTTTGTGTGCCAGCGCTGTGTCAGCGCTATGCGGGGCCAAAGCGGCCCTTCCCG 4320
4280 TATTTTGTGTGCCAGCGCTGTGTCAGCGCTATGCGGGGCCAAAGCGGCCCTTCCCG 4339
4321 CACGAGTATGTACGCGGGAACCGCGCATCGTGCCCTCAATTTCAATGACCGCGCGGTTC 4380
4340 CACGAGTATGTACGCGGGAACCGCGCATCGTGCCCTCAATTTCAATGACCGCGCGGTTC 4399
4381 CAGCATGCGCCCTTTCACAGGCGATCGCATGCGGGAAGTCCATGATGAGTCTCGTGAAGCTG 4440
4400 CAGCATGCGCCCTTTCACAGGCGATCGCATGCGGGAAGTCCATGATGAGTCTCGTGAAGCTG 4459
4441 ATGGGGGCGCGGGCGGGTCCCTCTGTACGACCGGAACACGTCACAGGGGCTCTCGTCC 4500
4460 ATGGGGGCGCGGGCGGGTCCCTCTGTACGACCGGAACACGTCACAGGGGCTCTCGTCC 4519
4501 AGCAGCTGTCGACGACGAAAGGCGCATCGTGTACCGCGCGATCTCTGAAACCGCGCGCTCC 4560
4520 AGCAGCTGTCGACGACGAAAGGCGCATCGTGTACCGCGCGATCTCTGAAACCGCGCGCTCC 4579
4561 CCGGCCACGGAACCCCTCCCTGTACAAATGACATGTTCTTACTCTTCAAAATTTCCGGCC 4620
4580 CCGGCCACGGAACCCCTCCCTGTACAAATGACATGTTCTTACTCTTCAAAATTTCCGGCC 4639
4621 ACTCGGAGACCGTACAGGCGCTACATCATTCAGGAATGGCGCCCGCCAGCAGCGCTCCG 4680
4640 ACTCGGAGACCGTACAGGCGCTACATCATTCAGGAATGGCGCCCGCCAGCAGCGCTCCG 4699
4681 AGCACCGAGCTGTGACAGCGACTACAGCGCCAGCGCTGGAAGGCGCAGCAAGTACTAC 4740
4700 AGCACCGAGCTGTGACAGCGACTACAGCGCCAGCGCTGGAAGGCGCAGCAAGTACTAC 4759
4741 CTGATTTGAACTCGGACTCAGACCCCTATCCACCCCGCCAGCAGCGCCAGCAGTAC 4800
4760 CTGATTTGAACTCGGACTCAGACCCCTATCCACCCCGCCAGCAGCGCCAGCAGTAC 4819
4801 CTGCGCGGAGGACAGCTGCGCCCTCGCGCCCGCCAGCAGGAGGAGTACTTCCATCTC 4860
4820 CTGTCGGGAGGAGACAGTGCCTCGCGCCCTCGCGCCCGCCAGCAGGAGGAGTACTTCCATCTC 4879
4861 TTCCCGCCCTCTCGTCCCTGCAAGGACTCATCTGACCTCGCGCGCGGCGCTCTGGC 4920
4880 TTCCCGCCCTCTCGTCCCTGCAAGGACTCATCTGACCTCGCGCGCGGCGCTCTGGC 4939
4921 TTCTCTGCGCCCTGTAATAGTTTAAATATGAACAAAGAAATAATATTTATCAT 4980
4940 TTCTCTGCGCCCTGTAATAGTTTAAATATGAACAAAGAAATAATATTTATCAT 4999
4981 TTAATAATAATAATAATAGTTTAAATAATGAAATGAGAAATGTGAATGTGATGGGGTG 5040
5000 TTAATAATAATAATAATAGTTTAAATAATGAAATGAGAAATGTGAATGTGATGGGGTG 5059
5041 GGCAGGCGTGGAGAACTTTGTA 5063
5060 GGCAGGCGTGGAGAACTTTGTA 5082

Db 1760 CAGCGCCGAGCATCGAGCGGTGCAAGAGTCAAGGCGAGCGGAGCGTCACTATTGAC 1819
Qy 1801 CAGCTGCCCGACCTGATCGGGCTCAAGCTGTGATGTGGCCAGGTCGTGGACCAAC 1860
Db 1820 CAGCTGCCCGACCTGATCGGGCTCAAGAGTGTGAATGTGGCCAGGTCGTGGAAACCAAC 1879
Qy 1861 CCGTGTGGCGGACAGGAACGGGGGGTGCAGCCACCTGTGCTTCTTCAACACCCACGCAACC 1920
Db 1880 CCGTGTGGCGGACAGGAACGGGGGGTGCAGCCACCTGTGCTTCTTCAACACCCACGCAACC 1939
Qy 1921 CCGTGTGGCTGCCCATCGGCTGTGAGCTGTGAGTGAATGAAGACCTGTGATGTGCTT 1980
Db 1940 CCGTGTGGCTGCCCATCGGCTGTGAGCTGTGAGTGAATGAAGACCTGTGATGTGCTT 1999
Qy 1981 GAGGCTTCTTGTGCTTCAACAGCAGAGCGCGCCATCCACAGATCTCCCTCGAGACCAAT 2040
Db 2000 GAGGCTTCTTGTGCTTCAACAGCAGAGCGCGCCATCCACAGATCTCCCTCGAGACCAAT 2059
Qy 2041 AACACGAGCTGGCCATCCCGCTCACGGGGCGTCAAGGAGGCTCAGCCCTCGACTTTGAT 2100
Db 2060 AACACGAGCTGGCCATCCCGCTCACGGGGCGTCAAGGAGGCTCAGCCCTCGACTTTGAT 2119
Qy 2101 GTGTCCAAACAACCATCTA CTGGA CAGAGTCAAGCTTGAAGACCATCAGCCGCGCTTC 2160
Db 2120 GTGTCCAAACAACCATCTA CTGGA CAGAGTCAAGCTTGAAGACCATCAGCCGCGCTTC 2179
Qy 2161 ATGAACGGAGCTGGTGGAGCAGTGTGTGAGTTGGCTTACTACCCGAGGGCATG 2220
Db 2180 ATGAACGGAGCTGGTGGAGCAGTGTGTGAGTTGGCTTACTACCCGAGGGCATG 2239
Qy 2221 GCCGTGACTGGATGGGCAAGAACTCTACTGGGCGGACACTGGAGCCAAACAGAACTGAA 2280
Db 2240 GCCGTGACTGGATGGGCAAGAACTCTACTGGGCGGACACTGGAGCCAAACAGAACTGAA 2299
Qy 2281 GTGGCGGGCTGGAGCGGAGTTCGGCAAGTCTCTGTGTGGAGGAGTTCGGAACAACCG 2340
Db 2300 GTGGCGGGCTGGAGCGGAGTTCGGCAAGTCTCTGTGTGGAGGAGTTCGGAACAACCG 2359
Qy 2341 AGTGTGCTGGCCCTGGATCCCAACAAAGGCTACATCTACTGGACCGAGTGGGGCGGCAAG 2400
Db 2360 AGTGTGCTGGCCCTGGATCCCAACAAAGGCTACATCTACTGGACCGAGTGGGGCGGCAAG 2419
Qy 2401 CCGAGGATCTGTGGGGCTTCATGAGCGGGACCAACTGCATGACGCTGGTGGACAAAGTG 2460
Db 2420 CCGAGGATCTGTGGGGCTTCATGAGCGGGACCAACTGCATGACGCTGGTGGACAAAGTG 2479
Qy 2461 GGCGGGCCAAACGACTTCAACATTGACTACGCTGACAGCGCTCTACTGGAACCGACTG 2520
Db 2480 GGCGGGCCAAACGACTTCAACATTGACTACGCTGACAGCGCTCTACTGGAACCGACTG 2539
Qy 2521 GACACCAATGATCGAGTCTGTCNAACATGCTGGGTGAGGAGCGGTGCTGATTCGCGAC 2580
Db 2540 GACACCAATGATCGAGTCTGTCNAACATGCTGGGTGAGGAGCGGTGCTGATTCGCGAC 2599
Qy 2581 GATCTCCGCGACCGCTTCCGCTGTGACGAGTACAGCGATTAATATCTACTGACAGACTGG 2640
Db 2600 GATCTCCGCGACCGCTTCCGCTGTGACGAGTACAGCGATTAATATCTACTGACAGACTGG 2659
Qy 2641 AATCTGCA CAGCATGAGCGGGCGGACAAAGACTAGCGGCGGAAACCGCACCTCATCCAG 2700
Db 2660 AATCTGCA CAGCATGAGCGGGCGGACAAAGACTAGCGGCGGAAACCGCACCTCATCCAG 2719
Qy 2701 GGCCACTGGAATCTGTGATGAGCATCTGTGTGTTCCACTCTCCCGCAGAGTGGCTC 2760
Db 2720 GGCCACTGGAATCTGTGATGAGCATCTGTGTGTTCCACTCTCCCGCAGAGTGGCTC 2779
Qy 2761 AATGACTGTATGCAACAACGGGCGAGTGGGAGCTGTGCTTGGCCATCCCGGGCGG 2820
Db 2780 AATGACTGTATGCAACAACGGGCGAGTGGGAGCTGTGCTTGGCCATCCCGGGCGG 2839
Qy 2821 CACCGCTCGGCTGGCTTCACTTACACCTTGAACCCAGCAGCCGCAACTGACGCGCG 2880

Db 2840 CACCGCTCGGCTGGCTGCTCACACTACACCTTGAACCCAGCAGCCGCAACTGACGCGG 2899
Qy 2881 CCCACCACTTCTTGTGCTTTCAGCCAGAAATCTGCCATCAGTCCGATCATCCGACGAC 2940
Db 2900 CCCACCACTTCTTGTGCTTTCAGCCAGAAATCTGCCATCAGTCCGATCATCCGACGAC 2959
Qy 2941 CAGCACAGCCCGGATCTCATCTCTGCCCCTGATGGACTGAGGAACTCAAGACCATCGAC 3000
Db 2960 CAGCACAGCCCGGATCTCATCTCTGCCCCTGATGGACTGAGGAACTCAAGACCATCGAC 3019
Qy 3001 TATGACCACTGTGACCAAGTTTCACTA CTGGGTGAGTGGGCGCAGAA CATAAGGAGCC 3060
Db 3020 TATGACCACTGTGACCAAGTTTCACTA CTGGGTGAGTGGGCGCAGAA CATAAGGAGCC 3079
Qy 3061 AAGGACGAGCGGAGCCAGCCCTTGTGTTTGTAGCTCTCTGAGCCAGGCGCAAAACCCAGAC 3120
Db 3080 AAGGACGAGCGGAGCCAGCCCTTGTGTTTGTAGCTCTCTGAGCCAGGCGCAAAACCCAGAC 3139
Qy 3121 AGGACGAGCCCAACGACCTCAGCATCGACATCTTACAGCCGGA CACTGTTCTGGACGTGCGAG 3180
Db 3140 AGGACGAGCCCAACGACCTCAGCATCGACATCTTACAGCCGGA CACTGTTCTGGACGTGCGAG 3199
Qy 3181 GCCACCAATACCATCAACGTCCACAGGCTGAGCGGGGAGCCATG GGGGTGTGCTGGT 3240
Db 3200 GCCACCAATACCATCAACGTCCACAGGCTGAGCGGGGAGCCATG GGGGTGTGCTGGT 3259
Qy 3241 GGGGACCGGACAGCCAGGCGCATCGTCTGCTCAACCGGAGCGAGG GTACTGTTACTTC 3300
Db 3260 GGGGACCGGACAGCCAGGCGCATCGTCTGCTCAACCGGAGCGAGG GTACTGTTACTTC 3319
Qy 3301 ACCAACATGACAGGACCGGGCAGCCAAAGATCGAA CGCAGCCCTTGGACGGCACCCAGGCGC 3360
Db 3320 ACCAACATGACAGGACCGGGCAGCCAAAGATCGAA CGCAGCCCTTGGACGGCACCCAGGCGC 3379
Qy 3361 GAGTCTCTTTCACACACCGGCTCATTCGCCCTGTGGCCCTGTGG TGGTGGACAAACACTG 3420
Db 3380 GAGTCTCTTTCACACACCGGCTCATTCGCCCTGTGGCCCTGTGG TGGTGGACAAACACTG 3439
Qy 3421 GGCAGGCTGTTCTGGGTGGACCGGACCTGAAGCGCATTTGAGAG CTGTGACCTGTGACGG 3480
Db 3440 GGCAGGCTGTTCTGGGTGGACCGGACCTGAAGCGCATTTGAGAG CTGTGACCTGTGACGG 3499
Qy 3481 GCCAACCGCTGACCTCTGGAGAGCGCCAAACATCTGTGACGCTCT TGGGCTGACCATCTCT 3540
Db 3500 GCCAACCGCTGACCTCTGGAGAGCGCCAAACATCTGTGACGCTCT TGGGCTGACCATCTCT 3559
Qy 3541 GGCAGGATCTCTA CTGGATCGACCGCAGCAGCAGATGATCGAG CTGTGGAGAGAC 3600
Db 3560 GGCAGGATCTCTA CTGGATCGACCGCAGCAGCAGATGATCGAG CTGTGGAGAGAC 3619
Qy 3601 ACCGGGACAGCGGACTCGCATCCAGGCGCTGTGCGCCACCTCAGT GGCATCCATGCA 3660
Db 3620 ACCGGGACAGCGGACTCGCATCCAGGCGCTGTGCGCCACCTCAGT GGCATCCATGCA 3679
Qy 3661 GTGAGGAAAGTCACTCTGTGGAGAGTCTTCTACGCCCA CCACTGTGCCCTGCAATGGTGGC 3720
Db 3680 GTGAGGAAAGTCACTCTGTGGAGAGTCTTCTACGCCCA CCACTGTGCCCTGCAATGGTGGC 3739
Qy 3721 TGCTCCCACTCTGTATTGTCAGAGGATGGGACACACACG GTGTCTATGCCCAGTCCAC 3780
Db 3740 TGCTCCCACTCTGTATTGTCAGAGGATGGGACACACACG GTGTCTATGCCCAGTCCAC 3799
Qy 3781 CTGCTGTCTTCGCAAGACCTGCTGACCTGTGAGAGCGCCCACTG CTCTCCCGGACGAG 3840
Db 3800 CTGCTGTCTTCGCAAGACCTGCTGACCTGTGAGAGCGCCCACTG CTCTCCCGGACGAG 3859
Qy 3841 TTTTCATGTGCA CAGGGGAGATCGACTGTATCCCGGGGCTTGGCG CTGTGAAGCGCTTT 3900
Db 3860 TTTTCATGTGCA CAGGGGAGATCGACTGTATCCCGGGGCTTGGCG CTGTGTGACGGCTTT 3919
Qy 3901 CCCGAGTGCATGATCAGAGCGACGAGAGGGGTGCCCCGCTGTGTG TCTCCGCGCCAGTTC 3960
Db 3920 CCCGAGTGCATGATCAGAGCGACGAGAGGGGTGCCCCGCTGTGTG TCTCCGCGCCAGTTC 3979

Db 500 GTGCTCTTCGAGGACTTTGACACCGCGAGGGCCATCGCTTTGGACCCCGCTCACGGG 559
Qy 541 TACATGTACTCGACAGACTGGGGTGTAGACGCCCGCGATTGAGCGGGCAGGGATGATGGC 600
Db 560 TACATGTACTCGACAGACTGGGGTGTAGACGCCCGCGATTGAGCGGGCAGGGATGATGGC 619
Qy 601 AGCACCCGGAAGATCATTTGTGGACTCGGACATTTACTTGGCCCAATGGACTGACCATCGAC 660
Db 620 AGCACCCGGAAGATCATTTGTGGACTCGGACATTTACTTGGCCCAATGGACTGACCATCGAC 679
Qy 661 CTGAGAGGACGAGAAGCTTACTTGGGCTGACGCCAAGCTTCAGCTTTCATCCACCGTGCCAAAC 720
Db 680 CTGAGAGGACGAGAAGCTTACTTGGGCTGACGCCAAGCTTCAGCTTTCATCCACCGTGCCAAAC 739
Qy 721 CTGAGAGGCTGTTCCGCGACAGAGGTGTGTGAGGGCAGCCTGACGCCACCCCTTCGCCCTG 780
Db 740 CTGAGAGGCTGTTCCGCGACAGAGGTGTGTGAGGGCAGCCTGACGCCACCCCTTCGCCCTG 799
Qy 781 ACGCTTCTCCGGGACACTCTGTACTTGGACAGACTGGCAGACCCCGCTCCATCCATGCCCTGC 840
Db 800 ACGCTTCTCCGGGACACTCTGTACTTGGACAGACTGGCAGACCCCGCTCCATCCATGCCCTGC 859
Qy 841 AACAAAGCGCACTTGGGGGGAAGAGGAGATCTCTGAGTGCCCTCTACTCAACCCATGGAC 900
Db 860 AACAAAGCGCACTTGGGGGGAAGAGGAGATCTCTGAGTGCCCTCTACTCAACCCATGGAC 919
Qy 901 ATCCAGGTGTGAGCAGGAGCGGCGAGCCTTTTCTTCCACACTCGTGTGTGAGGAGACAAT 960
Db 920 ATCCAGGTGTGAGCAGGAGCGGCGAGCCTTTTCTTCCACACTCGTGTGTGAGGAGACAAT 979
Qy 961 GCGCGCTGCTCCACCTGTGCTGCTGTCCCAAGCGAGCCTTTCTACACATGCGCCTGC 1020
Db 980 GCGCGCTGCTCCACCTGTGCTGCTGTCCCAAGCGAGCCTTTCTACACATGCGCCTGC 1039
Qy 1021 CCCACGGGTGTGAGCTGCAGGACAAACCGCAGGAGCGTGTAAAGGAGGAGCGCGAGGAGTG 1080
Db 1040 CCCACGGGTGTGAGCTGCAGGACAAACCGCAGGAGCGTGTAAAGGAGGAGCGCGAGGAGTG 1099
Qy 1081 CTGCTGTGCGCGCGGCGAGCGGACCTTACGAGAGATCTCGCTGTGAGACACCGCGGACTTCACC 1140
Db 1100 CTGCTGTGCGCGCGGCGAGCGGACCTTACGAGAGATCTCGCTGTGAGACACCGCGGACTTCACC 1159
Qy 1141 GACATCGTGTGAGTGTGAGCGACATCCGCGACGCCATTTGCCATCGACTAGACCCGCTA 1200
Db 1160 GACATCGTGTGAGTGTGAGCGACATCCGCGACGCCATTTGCCATCGACTAGACCCGCTA 1219
Qy 1201 GAGGGCTATGTCTACTGACAGATGACAGAGGTGCGGGCCATCCGCGAGGCGGTACCTGGAC 1260
Db 1220 GAGGGCTATGTCTACTGACAGATGACAGAGGTGCGGGCCATCCGCGAGGCGGTACCTGGAC 1279
Qy 1261 GGGTCTGGGGCGGACGCTCGTGTAAACACCGAGATCAACGACCCCGATGGCATCGCGTGC 1320
Db 1280 GGGTCTGGGGCGGACGCTCGTGTAAACACCGAGATCAACGACCCCGATGGCATCGCGTGC 1339
Qy 1321 GACTGGGTGGCCGAAACCTCTACTGAGCCGACGACCGGCGACCGCATCGAGGTGACG 1380
Db 1340 GACTGGGTGGCCGAAACCTCTACTGAGCCGACGACCGGCGACCGCATCGAGGTGACG 1399
Qy 1381 GCGCTCAACCGGACCTCCCGCAAGATCTCGTGTGTGAGGAGCCTTGGACAGCGCCCGAGGCC 1440
Db 1400 GCGCTCAACCGGACCTCCCGCAAGATCTCGTGTGTGAGGAGCCTTGGACAGCGCCCGAGGCC 1459
Qy 1441 ATCGCACTGACACCCCGTGTATGGGCTCATGTACTGGAACAGACTGGGGAGAGAACCTTAAA 1500
Db 1460 ATCGCACTGACACCCCGTGTATGGGCTCATGTACTGGAACAGACTGGGGAGAGAACCTTAAA 1519
Qy 1501 ATCGAGTGTCCAACTTGGATGGGCGAGGCGGTGTGTGTTCAATGCGCTCCCTCGG 1560
Db 1520 ATCGAGTGTCCAACTTGGATGGGCGAGGCGGTGTGTGTTCAATGCGCTCCCTCGG 1579
Qy 1561 TGGCCCAACGGCTTGGACCTTGCAGAGGGGGAAGCTTACTTGGGGGAGACGCCAAG 1620

Db 1580 TGGCCCAACGGCTTGGACCTTGCAGAGAGGGGAAGCTTACTTGGGGAGACGCCAAG 1639
Qy 1621 ACAGACAGATCGAGGTGATCAATGTTGATGGGACGAGAGGGCGGACCTCTCTCGAGGAC 1680
Db 1640 ACAGACAGATCGAGGTGATCAATGTTGATGGGACGAGAGGGCGGACCTCTCTCGAGGAC 1699
Qy 1681 AAGCTCCCGCACATTTTTCGGGTTTTCACGCTGCTGGGGGACTTTCATCTACTGAGTACTGG 1740
Db 1700 AAGCTCCCGCACATTTTTCGGGTTTTCACGCTGCTGGGGGACTTTCATCTACTGAGTACTGG 1759
Qy 1741 CAGGCGCGAGCATCGAGCGGGTGCACAAGGTCAAGGCGACCGCGGAGCGTCAATATTGAC 1800
Db 1760 CAGGCGCGAGCATCGAGCGGGTGCACAAGGTCAAGGCGACCGCGGAGCGTCAATATTGAC 1819
Qy 1801 CAGCTGCGCGACCTGATGGGGCTCAAAGCTGTGAAATGTGGCCAAAGGTGCTCGGAACCAAC 1860
Db 1820 CAGCTGCGCGACCTGATGGGGCTCAAAGCTGTGAAATGTGGCCAAAGGTGCTCGGAACCAAC 1879
Qy 1861 CCGTGTGCGGACAGGAAACGGGGGTGCAGCCACCTGTGCTTCTTTCACACCCCAAGCAACC 1920
Db 1880 CCGTGTGCGGACAGGAAACGGGGGTGCAGCCACCTGTGCTTCTTTCACACCCCAAGCAACC 1939
Qy 1921 CCGTGTGCGTGCCTCCCATCGGCTCGGAGTGTGTGATGACATGAAGACCTGTGATCGTGCCT 1980
Db 1940 CCGTGTGCGTGCCTCCCATCGGCTCGGAGTGTGTGATGACATGAAGACCTGTGATCGTGCCT 1999
Qy 1981 GAGGCTTCTTGGTCTTTCACGAGCAGAGCGGCCATCCACAGGATCTCCCTCGAGACCAAT 2040
Db 2000 GAGGCTTCTTGGTCTTTCACGAGCAGAGCGGCCATCCACAGGATCTCCCTCGAGACCAAT 2059
Qy 2041 AACAAACGAGCTGGGCATCCCGCTCAACGGGCTCAAGGAGGCTCAGGCCCTGGACTTTGAT 2100
Db 2060 AACAAACGAGCTGGGCATCCCGCTCAACGGGCTCAAGGAGGCTCAGGCCCTGGACTTTGAT 2119
Qy 2101 GTGTCCAAACCAATCTACTTGAACAGAGCTGAGCTGTGAAGACATCAGCGCGGCTTC 2160
Db 2120 GTGTCCAAACCAATCTACTTGAACAGAGCTGAGCTGTGAAGACATCAGCGCGGCTTC 2179
Qy 2161 ATGAAACGGGAGCTCGGTGGAGCAGCTGGTGGAGTTTGGCCCTTGACTACCCCGAGGGCATG 2220
Db 2180 ATGAAACGGGAGCTCGGTGGAGCAGCTGGTGGAGTTTGGCCCTTGACTACCCCGAGGGCATG 2239
Qy 2221 GCCCTTGACTGGATGGGCAAGAACTCTACTGGGCGGACACTGGGACCAACAGAAATCGAA 2280
Db 2240 GCCCTTGACTGGATGGGCAAGAACTCTACTGGGCGGACACTGGGACCAACAGAAATCGAA 2299
Qy 2281 GTGCGCGGTGTGAGCGGGAGTTCGGCAAGTCTCTGTGTGGAGGAGCTTGGACAAACCGG 2340
Db 2300 GTGCGCGGTGTGAGCGGGAGTTCGGCAAGTCTCTGTGTGGAGGAGCTTGGACAAACCGG 2359
Qy 2341 AGGTGCGTGGCCCTGGATCCCAACAGGGCTTACATCTACTGAGCCGAGTGGGGCGGCAAG 2400
Db 2360 AGGTGCGTGGCCCTGGATCCCAACAGGGCTTACATCTACTGAGCCGAGTGGGGCGGCAAG 2419
Qy 2401 CCGAGGATCGTGGCGGCTTTCATGAGCGGAGCAACCTGATGATGAGCTGGTGGACAGGCTG 2460
Db 2420 CCGAGGATCGTGGCGGCTTTCATGAGCGGAGCAACCTGATGATGAGCTGGTGGACAGGCTG 2479
Qy 2461 GCGCGGGCAACAGACTTTCATGAGCGGAGCAACCTGATGATGAGCTGGTGGACAGGCTG 2520
Db 2480 GCGCGGGCAACAGACTTTCATGAGCGGAGCAACCTGATGATGAGCTGGTGGACAGGCTG 2539
Qy 2521 GACACCAACATGATCGAGTGTGCAACATGCTGGGTGAGGAGCGGCTGATGATGCGGCAC 2580
Db 2540 GACACCAACATGATCGAGTGTGCAACATGCTGGGTGAGGAGCGGCTGATGATGCGGCAC 2599
Qy 2581 GATCTCCCGCAACCGCTTTCGCTTTCAGCAGTACAGGATTTATATCTACTTGGACAGACTGG 2640
Db 2600 GATCTCCCGCAACCGCTTTCGCTTTCAGCAGTACAGGATTTATATCTACTTGGACAGACTGG 2659
Qy 2641 AATCTGACAGCATTTGAGCGGGCGGACAAAGACTAGCGGCGGAAACCGCACCCCTCATCCAG 2700
Db 2660 AATCTGACAGCATTTGAGCGGGCGGACAAAGACTAGCGGCGGAAACCGCACCCCTCATCCAG 2719

QY 2701 GSCACCTGGACTTCTGTGATGACATCTGTGTGTTCCACTCTCTCCGCGCAGGATGGCCTC 2760
Db 2720 GSCACCTGGACTTCTGTGATGACATCTGTGTGTTCCACTCTCTCCGCGCAGGATGGCCTC 2779
QY 2761 AATGACTGTATGACAAACAAACGGGAGTGTGTGGCAGCTGTGTGCTTGGCCATCCCGCGCGC 2820
Db 2780 AATGACTGTATGACAAACAAACGGGAGTGTGTGGCAGCTGTGTGCTTGGCCATCCCGCGCGC 2839
QY 2821 CACCGCTGGGCTGGCCTTCACTATACACCTTGGACCCAGAGCGGCAATGTGACGCGC 2880
Db 2840 CACCGCTGGGCTGGCCTTCACTATACACCTTGGACCCAGAGCGGCAATGTGACGCGC 2899
QY 2881 CCCACACCTTCTGTCTGTTCAGCCAGAGAAATCTGCCATCAGTCGGATGATCCCGACGAC 2940
Db 2900 CCCACACCTTCTGTCTGTTCAGCCAGAGAAATCTGCCATCAGTCGGATGATCCCGACGAC 2959
QY 2941 CAGCAGCCCCGATCTCATCTGCCCCCTGCTGACTGAGGAACTGCAAGGCCATCGAC 3000
Db 2960 CAGCAGCCCCGATCTCATCTGCCCCCTGCTGACTGAGGAACTGCAAGGCCATCGAC 3019
QY 3001 TATGACCACTGGACAAAGTTTCACTACTGGGTGGATGGCGGCAGAAATCAAGCGAGCC 3060
Db 3020 TATGACCACTGGACAAAGTTTCACTACTGGGTGGATGGCGGCAGAAATCAAGCGAGCC 3079
QY 3061 AAGGACGAGCGGACCCAGCCCTTTGTTTGTGACCTCTCTGAGCCAGGCGCAAAACCCAGAC 3120
Db 3080 AAGGACGAGCGGACCCAGCCCTTTGTTTGTGACCTCTCTGAGCCAGGCGCAAAACCCAGAC 3139
QY 3121 AAGGACGCCCCAGACCTCAGCATCGACATCTACAGCCGGAACAATGTTCTGGACGTGCGAG 3180
Db 3140 AAGGACGCCCCAGACCTCAGCATCGACATCTACAGCCGGAACAATGTTCTGGACGTGCGAG 3199
QY 3181 GCCACCAATACATCAAGCTTCAAGGCTGAGCGGGGAAGCCATGGGGTGGTGTGCGT 3240
Db 3200 GCCACCAATACATCAAGCTTCAAGGCTGAGCGGGGAAGCCATGGGGTGGTGTGCGT 3259
QY 3241 GGGGACCGCGACAAGCCAGGSCCATCTGCTGTAACGGGAGCGAGGGTACTCTGTACTTC 3300
Db 3260 GGGGACCGCGACAAGCCAGGSCCATCTGCTGTAACGGGAGCGAGGGTACTCTGTACTTC 3319
QY 3301 ACCAACTGCGAGGACCGGGGAGCCAAAGATCGAAACGCGAGCCCTGGACGCGACCGAGCGC 3360
Db 3320 ACCAACTGCGAGGACCGGGGAGCCAAAGATCGAAACGCGAGCCCTGGACGCGACCGAGCGC 3379
QY 3361 GAGTCTCTTTACACACCGGCTCATCCGCCCTGTGGCCCTGTGGTGGACAAACACTG 3420
Db 3380 GAGTCTCTTTACACACCGGCTCATCCGCCCTGTGGCCCTGTGGTGGACAAACACTG 3439
QY 3421 GCGAAGCTGTTCTGGGTGGACGCGACCTGAAGCGCATTTGAGAGCTGTGACCTGTGAGGG 3480
Db 3440 GCGAAGCTGTTCTGGGTGGACGCGACCTGAAGCGCATTTGAGAGCTGTGACCTGTGAGGG 3499
QY 3481 GCCAACCGCCTGACCTGGAGGACGCCAAACATCTGTGAGCCCTCTGGGCTGACCATCCTT 3540
Db 3500 GCCAACCGCCTGACCTGGAGGAGCCAAACATCTGTGAGCCCTCTGGGCTGACCATCCTT 3559
QY 3541 GGCAAGCATCTTACTGTATGACCGGCCAGAGCAGATGATCGAGCGTGTGGAGAGCC 3600
Db 3560 GGCAAGCATCTTACTGTATGACCGGCCAGAGCAGATGATCGAGCGTGTGGAGAGCC 3619
QY 3601 ACCGGGACAAGCGGACTCGCATCCAGGGCGTGTGCGCCACCTCACTGGCATCCATGCA 3660
Db 3620 ACCGGGACAAGCGGACTCGCATCCAGGGCGTGTGCGCCACCTCACTGGCATCCATGCA 3679
QY 3661 GTGAGGAAAGTCAAGCTGGAGGAGTTCTCAGCCCAACCAATGTGCCCGTGAATGGTGGC 3720
Db 3680 GTGAGGAAAGTCAAGCTGGAGGAGTTCTCAGCCCAACCAATGTGCCCGTGAATGGTGGC 3739
QY 3721 TGCTCCCAACATCTGTATTGCAAGGGTGAATGGGACCAACCGGTGCTATGCCCGAGTCCAC 3780
Db 3740 TGCTCCCAACATCTGTATTGCAAGGGTGAATGGGACCAACCGGTGCTATGCCCGAGTCCAC 3799

QY 3781 CTCGTGCTCTCGCAGAACCTGCTGACCTGTGTGAGAGCGCGCCACCTGCTCCCGGACGAC 3840
Db 3800 CTCGTGCTCTCGCAGAACCTGCTGACCTGTGTGAGAGCGCGCCACCTGCTCCCGGACGAC 3859
QY 3841 TTTGCACTGTGCGACAGGGGAGATCGACTGTATCTCCCGGGGCTGTGGCGCTGTGACGGCTTT 3900
Db 3860 TTTGCACTGTGCGACAGGGGAGATCGACTGTATCTCCCGGGGCTGTGGCGCTGTGACGGCTTT 3919
QY 3901 CCCAGATGCGATGACAGAGCGACAGAGAGGGCTGCCCCGTGTGTGCTCCGCGCCGAGTTC 3960
Db 3920 CCCAGATGCGATGACAGAGCGACAGAGAGGGCTGCCCCGTGTGTGCTCCGCGCCGAGTTC 3979
QY 3961 CCCTGCGCGCGGGTCAAGTGTGTGACCTGCGCTGCGCTGCGAGCGGAGGAGAGTGT 4020
Db 3980 CCCTGCGCGCGGGTCAAGTGTGTGACCTGCGCTGCGCTGCGAGCGGAGGAGAGTGT 4039
QY 4021 CAGGACCGCTCAGACGAGGGCGACTGTGACCGCCATCTGCTTGCCTGCCCAACACAGTTCGGTGT 4080
Db 4040 CAGGACCGCTCAGACGAGGGTGTGACTGTGACCGCATCTGCTTGCCTGCCCAACACAGTTCGGTGT 4099
QY 4081 GCGAGCGGCGAGTGTGCTCTCATCAAAACAGAGTGTGACTCTTCCCGGACTGTATCGAC 4140
Db 4100 GCGAGCGGCGAGTGTGCTCTCATCAAAACAGAGTGTGACTCTTCCCGGACTGTATCGAC 4159
QY 4141 GGCCTCCGACGAGCTCATGTGTGAAATCACCAAGCGGCTCAGACGACGAGCCGCGCCAC 4200
Db 4160 GGCCTCCGACGAGCTCATGTGTGAAATCACCAAGCGGCTCAGACGACGAGCCGCGCCAC 4219
QY 4201 AGCAGTGCATFCGGGCGCGTCAATTTGGCATCATCTCTCTCTCTTCTGTCATGGGTGGTGC 4260
Db 4220 AGCAGTGCATFCGGGCGCGTCAATTTGGCATCATCTCTCTCTCTCTTCTGTCATGGGTGGTGC 4279
QY 4261 TATTTTGTGTGCGACGCGGTGGTGTGCGAGGCTATGTGGGGGCGAAGCGGCGCTTCCG 4320
Db 4280 TATTTTGTGTGCGACGCGGTGGTGTGCGAGGCTATGTGGGGGCGAAGCGGCGCTTCCG 4339
QY 4321 CACGAGTATGTACGCGGACCCCGACGTCGCTCAATTTATAGCCCGGGGGTTC 4380
Db 4340 CACGAGTATGTACGCGGACCCCGACGTCGCTCAATTTATAGCCCGGGGGTTC 4399
QY 4381 CAGCATGGCCCTTTCACAGGATCGCATGCGGAAAGTCCATGATGAGCTCCGTGAGCGTG 4440
Db 4400 CAGCATGGCCCTTTCACAGGATCGCATGCGGAAAGTCCATGATGAGCTCCGTGAGCGTG 4459
QY 4441 ATGGGGGCGGGGCGGGTGCCTCTGACGCGGAAACACGTCACAGGGGCTCTGTC 4500
Db 4460 ATGGGGGCGGGGCGGGTGCCTCTGACGCGGAAACACGTCACAGGGGCTCTGTC 4519
QY 4501 AGCAGCTCGTCCAGACAGAGGCGACGCTGTACCGCGCGATCTCTGAACCCCGCCCTTC 4560
Db 4520 AGCAGCTCGTCCAGACAGAGGCGACGCTGTACCGCGCGATCTCTGAACCCCGCCCTTC 4579
QY 4561 CCGGCCACGAGCCCTCCCTGTACAAATGGACATGTTTACTCTTCAAAATTCGCGCC 4620
Db 4580 CCGGCCACGAGCCCTCCCTGTACAAATGGACATGTTTACTCTTCAAAATTCGCGCC 4639
QY 4621 ACTCGGAGCCGTACAGCCCTTACATTCGAGGAATGGCGCCCGCCGACGAGCCCTTC 4680
Db 4640 ACTCGGAGCCGTACAGCCCTTACATTCGAGGAATGGCGCCCGCCGACGAGCCCTTC 4699
QY 4681 AGCACCGAGCTGTGTGACGAGCTACAGCGCGCAGCGCTGGAAGGCGCAGCACTATAC 4740
Db 4700 AGCACCGAGCTGTGTGACGAGCTACAGCGCGCAGCGCTGGAAGGCGCAGCACTATAC 4759
QY 4741 CTGGAATTTGAATCTGGAATCTAGACCCCTATCCACCCCAACCGCCCGCCACGAGCTAC 4800
Db 4760 CTGGAATTTGAATCTGGAATCTAGACCCCTATCCACCCCAACCGCCCGCCACGAGCTAC 4819
QY 4801 CTGTGCGGAGGAGGAGCTGCGCCCTTCCCGCGCCACCGAGAGGAGTACTTCCATCTC 4860
Db 4820 CTGTGCGGAGGAGGAGCTGCGCCCTTCCCGCGCCACCGAGAGGAGTACTTCCATCTC 4879
QY 4861 TTCCCGCCCTCTCGTCTCCCTCTGACGAGCTCATCTGACCTCGGCGGGGCGACTCTGGC 4920

||||| 1489 ATCGACTGCACCCCGTGTATGGGCTCATGTACTGACAGACTGGGGAGAGAACCTTAA 1548
QY ||||| 1501 ATCGAGTGTGCAACTTGGATGGGAGGAGCGCGTGTGCTGGTCAATGCCCTCCCTCGGG 1560
Db ||||| 1549 ATCGAGTGTGCAACTTGGATGGGAGGAGCGCGTGTGCTGGTCAATGCCCTCCCTCGGG 1608
QY ||||| 1561 TGGCCCAACGGCTGGCCCTGACCTGACGAGGGGAGCTTACTTGGGGAGAGCCCAAG 1620
Db ||||| 1609 TGGCCCAACGGCTGGCCCTGACCTGACGAGGGGAGCTTACTTGGGGAGAGCCCAAG 1668
QY ||||| 1621 ACAGACAAGATCGAGGTGATCAATGTGATGAGGACGAAAGAGCGGACCTCTCTGAGGAC 1680
Db ||||| 1669 ACAGACAAGATCGAGGTGATCAATGTGATGAGGACGAAAGAGCGGACCTCTCTGAGGAC 1728
QY ||||| 1681 AAGCTCCGCGCAATTTTGGGTTTCACTGCTGGGGGACTTCACTACTGGAAGTACTGG 1740
Db ||||| 1729 AAGCTCCGCGCAATTTTGGGTTTCACTGCTGGGGGACTTCACTACTGGAAGTACTGG 1788
QY ||||| 1741 CAGCGCGGAGATCGAGGGGTGCAAGGTCAAGGTCAAGCGGAGCGGAGCGTCAATTTGAC 1800
Db ||||| 1789 CAGCGCGGAGATCGAGGGGTGCAAGGTCAAGGTCAAGCGGAGCGGAGCGTCAATTTGAC 1848
QY ||||| 1801 CAGCTGCCCGCACTGATGGGGCTCAAAGCTGTGAATGTGGCCAAAGGTGTCGAAACCAAC 1860
Db ||||| 1849 CAGCTGCCCGCACTGATGGGGCTCAAAGCTGTGAATGTGGCCAAAGGTGTCGAAACCAAC 1908
QY ||||| 1861 CGGTGTGGGACAGAAACGGGGGTGACAGCCACTGTGCTTTTCAACCCCAAGCAAC 1920
Db ||||| 1909 CGGTGTGGGACAGAAACGGGGGTGACAGCCACTGTGCTTTTCAACCCCAAGCAAC 1968
QY ||||| 1921 CGGTGTGGCTGCCCATCGGCTGAGCTGCTGAGTGAATGATGAGACCTGATCGTGCT 1980
Db ||||| 1969 CGGTGTGGCTGCCCATCGGCTGAGCTGCTGAGTGAATGATGAGACCTGATCGTGCT 2028
QY ||||| 1981 GAGGCTTTCTTGGTCTTCAACAGAGAGCGCCATCCACAGATCTCCCTCGAGACCAAT 2040
Db ||||| 2029 GAGGCTTTCTTGGTCTTCAACAGAGAGCGCCATCCACAGATCTCCCTCGAGACCAAT 2088
QY ||||| 2041 AACAAAGAGTGGCCATCCCGTCAACGGGGTCAAGGAGGCTCAGCCCTGAGCTTTGAT 2100
Db ||||| 2089 AACAAAGAGTGGCCATCCCGTCAACGGGGTCAAGGAGGCTCAGCCCTGAGCTTTGAT 2148
QY ||||| 2101 GTGTCCAAACCAACATCTACTGACAGACCTCAGCTGAGACCAATCAGCGCGCTTC 2160
Db ||||| 2149 GTGTCCAAACCAACATCTACTGACAGACCTCAGCTGAGACCAATCAGCGCGCTTC 2208
QY ||||| 2161 ATGAACGGGAGCTCGGTGGAGCACTGTGTGGAGTTTGGCTTGAATACCCCGAGGGCATG 2220
Db ||||| 2209 ATGAACGGGAGCTCGGTGGAGCACTGTGTGGAGTTTGGCTTGAATACCCCGAGGGCATG 2268
QY ||||| 2221 GCGGTTGATCGGAGCAAGAACTTACTGAGGCGGACACTGGGACCAACAGAAATCGAA 2280
Db ||||| 2269 GCGGTTGATCGGAGCAAGAACTTACTGAGGCGGACACTGGGAGCAACAGAAATCGAA 2328
QY ||||| 2281 GTGGCGGGCTGAGCGGCGAGTTCGGCAAGTCTCTGAGTGGAGGACTTGGACAAACCG 2340
Db ||||| 2329 GTGGCGGGCTGAGCGGCGAGTTCGGCAAGTCTCTGAGTGGAGGACTTGGACAAACCG 2388
QY ||||| 2341 AGGTGCTGGCCCTGGATCCCAACCAAGGGCTTACATCTACTGAGCCGAGTGGGGGCGCAAG 2400
Db ||||| 2389 AGGTGCTGGCCCTGGATCCCAACCAAGGGCTTACATCTACTGAGCCGAGTGGGGGCGCAAG 2448
QY ||||| 2401 CCGAGGATCGGGGCTTTCATGAGCGGAGCAACACTGATGACGCTGGTGGACAGGTTG 2460
Db ||||| 2449 CCGAGGATCGGGGCTTTCATGAGCGGAGCAACACTGATGACGCTGGTGGACAGGTTG 2508
QY ||||| 2461 GCGCGGGCAACGACCTCACCATTGACTACGCTGACCAAGCGCTTACTTGAACGACCTG 2520
Db ||||| 2509 GCGCGGGCAACGACCTCACCATTGACTACGCTGACCAAGCGCTTACTTGAACGACCTG 2568
QY ||||| 2521 GACACCAACATGATCGAGTCTGCAACATGCTGGGTGAGGAGCGGGTGTGATTCGCGAC 2580

Db ||||| 2569 GACACCAACATGATCGAGTCTGCTCAACATGCTGGGTGAGGAGCGGGTGTGATTCGCGAC 2628
QY ||||| 2581 GATCTCCGCGCACCCGTTCTGGTCTGACGAGTACGAGCAATTATCTACTGACAGACTGG 2640
Db ||||| 2629 GATCTCCGCGCACCCGTTCTGGTCTGACGAGTACGAGCAATTATCTACTGACAGACTGG 2688
QY ||||| 2641 AATCTGACAGCATTTGAGCGGCGGACAAAGACTAGCGGCGGAAACCGACCTCATCCAG 2700
Db ||||| 2689 AATCTGACAGCATTTGAGCGGCGGACAAAGACTAGCGGCGGAAACCGACCTCATCCAG 2748
QY ||||| 2701 GGCACCTTGGACTTCTGATGAGCATCTCTGTGTTCACATCTCCCGCAGGATGGCCTC 2760
Db ||||| 2749 GGCACCTTGGACTTCTGATGAGCATCTCTGTGTTCACATCTCCCGCAGGATGGCCTC 2808
QY ||||| 2761 AATGACTGTATGTCAACAAACGGGCACTGTGGGAGCTGTGCTTGGCCATCCCCGGCGGC 2820
Db ||||| 2809 AATGACTGTATGTCAACAAACGGGCACTGTGGGAGCTGTGCTTGGCCATCCCCGGCGGC 2868
QY ||||| 2821 CACGCTCGGCTGGGCTTCACTACACCTTGACCCAGAGCGGCACTGACAGCCG 2880
Db ||||| 2869 CACGCTCGGCTGGGCTTCACTACACCTTGACCCAGAGCGGCACTGACAGCCG 2928
QY ||||| 2881 CCCACCACTTCTGTGCTTTCAGCCAGAAATCTGCATCAGTCTGAGTATCCCGAGCGAC 2940
Db ||||| 2929 CCCACCACTTCTGTGCTTTCAGCCAGAAATCTGCATCAGTCTGAGTATCCCGAGCGAC 2988
QY ||||| 2941 CAGCACAGCCCGGATCTCATCTCTCCCTGATGAGTCTGAGGAAACGTCACAAAGCCATCGAC 3000
Db ||||| 2989 CAGCACAGCCCGGATCTCATCTCTCCCTGATGAGTCTGAGGAAACGTCACAAAGCCATCGAC 3048
QY ||||| 3001 TATGACCCACTGGACAAAGTTTCATCTACTGGGTGATGGGCGCCAGAACATCAAGCGAGCC 3060
Db ||||| 3049 TATGACCCACTGGACAAAGTTTCATCTACTGGGTGATGGGCGCCAGAACATCAAGCGAGCC 3108
QY ||||| 3061 AAGACAGAGGAGCCAGCCCTTTGTTTGTACCTCTCTGAGCCAGGCGCAAAACCCAGAC 3120
Db ||||| 3109 AAGACAGAGGAGCCAGCCCTTTGTTTGTACCTCTCTGAGCCAGGCGCAAAACCCAGAC 3168
QY ||||| 3121 AGGACAGCCCAAGACTCTGAGCATCTGACATCTGACGCGGACACTGTCTTCTGACGCTGCGAG 3180
Db ||||| 3169 AGGACAGCCCAAGACTCTGAGCATCTGACATCTGACGCGGACACTGTCTTCTGACGCTGCGAG 3228
QY ||||| 3181 GGCACCAATACCAATCAAGCTCCAGGCTGAGCGGGGAGCCATGGGGGTGGTCTCGCT 3240
Db ||||| 3229 GGCACCAATACCAATCAAGCTCCAGGCTGAGCGGGGAGCCATGGGGGTGGTCTCGCT 3288
QY ||||| 3241 GGGGACCGGACAGCCAGGCGCATCTGCTCAACGCGAGCGAGGCTACTCTGACTTC 3300
Db ||||| 3289 GGGGACCGGAGCAAGCCAGGCGCATCTGCTCAACGCGAGCGAGGCTACTCTGACTTC 3348
QY ||||| 3301 ACCAAATGACAGGAGCCGGGAGCGCAAGATCGAAGCGGAGCCCTTGGAGCGGACCGAGCGC 3360
Db ||||| 3349 ACCAAATGACAGGAGCCGGGAGCGCAAGATCGAAGCGGAGCCCTTGGAGCGGACCGAGCGC 3408
QY ||||| 3361 GAGTCTCTTTCACCCCGGCTCATCCGCTGCTGGCCCTGCTGGTGGGAGCAACACACTG 3420
Db ||||| 3409 GAGTCTCTTTCACCCCGGCTCATCCGCTGCTGGCCCTGCTGGTGGGAGCAACACACTG 3468
QY ||||| 3421 GGCAGAGTCTTCTGGGTGGAGCGGAGCTTGAAGCGCATGAGAGCTGTGACTCTGACGGG 3480
Db ||||| 3469 GGCAGAGTCTTCTGGGTGGAGCGGAGCTTGAAGCGCATGAGAGCTGTGACTCTGTCAGGG 3528
QY ||||| 3481 GGCAGCGCTGACCTTGGAGGAGCGCAACATCTGTGAGCGCTTGGGCTGACCATCTCTT 3540
Db ||||| 3529 GGCAGCGCTGACCTTGGAGGAGCGCAACATCTGTGAGCGCTTGGGCTGACCATCTCTT 3588
QY ||||| 3541 GGCAGGATCTTCTACTGATCGACCGGACGAGCATGATCGAGCGCTTGGAGAGGAC 3600
Db ||||| 3589 GGCAGGATCTTCTACTGATCGACCGGACGAGCATGATCGAGCGCTTGGAGAGGAC 3648
QY ||||| 3601 ACCGGGAGCAAGCGGACTCTGCATCCAGGGCGGTGTGCGCCACCTCTACTTGGCATCCATGCA 3660
Db ||||| 3649 ACCGGGAGCAAGCGGACTCTGCATCCAGGGCGGTGTGCGCCACCTCTACTTGGCATCCATGCA 3708

Db 229 CGGCTGGTGGAGCCCGCGGAGTCAAGCTGGAGTCCACCATCGTGGTCAAGCGGCTGGAG 288
QY 241 GATCGCGCGCAGTGGACTTCCAGTCTTCCAGGGAGCGGTCTACTGGACAGACGTGAGC 300
Db 289 GATCGCGCGCAGTGGACTTCCAGTCTTCCAGGGAGCGGTCTACTGGACAGACGTGAGC 348
QY 301 GAGGAGGCATCAAGCAGACCTACCTGAACAGACGCGGGGCGCGGTGAGAACGTGTC 360
Db 349 GAGGAGGCATCAAGCAGACCTACCTGAACAGACGCGGGGCGCGGTGAGAACGTGTC 408
QY 361 ATCTCCGGCTCGTCTCTCCGACCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCT 420
Db 409 ATCTCCGGCTCGTCTCTCCGACCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCT 468
QY 421 TGGACGGACTCAGAGACCAACCGCATCGAGGTGGCCAACTCAATGGCACAATCCCGGAAG 480
Db 469 TGGACGGACTCAGAGACCAACCGCATCGAGGTGGCCAACTCAATGGCACAATCCCGGAAG 528
QY 481 GTGCTCTTCTGGCAGGACTTGAACAGCTTAGGGCCATCGCTTGGACCCGCTCACGGG 540
Db 529 GTGCTCTTCTGGCAGGACTTGAACAGCTTAGGGCCATCGCTTGGACCCGCTCACGGG 588
QY 541 TACATGTACTGACAGACTGGGGTGAACCGCCCGGATTGAGCGGGCAGGATGATG3C 600
Db 589 TACATGTACTGACAGACTGGGGTGAACCGCCCGGATTGAGCGGGCAGGATGATG3C 648
QY 601 AGCACCCGGAAGATCATTTGTGACTCGGACATTTTACTGGCCCAATGGACTGACCATCGAC 660
Db 649 AGCACCCGGAAGATCATTTGTGACTCGGACATTTTACTGGCCCAATGGACTGACCATCGAC 708
QY 661 CTGAGGAGCAGAGCTCTACTGGGCTGAGCGCCCGGATTGAGCGGGCAGGATGATG3C 720
Db 709 CTGAGGAGCAGAGCTCTACTGGGCTGAGCGCCCGGATTGAGCGGGCAGGATGATG3C 768
QY 721 CTGAGCGGCTCGTTCGGCAGAGGTGGTGGAGGCGCCTGAGCGACCCCTTGGCCCTG 780
Db 769 CTGAGCGGCTCGTTCGGCAGAGGTGGTGGAGGCGCCTGAGCGACCCCTTGGCCCTG 828
QY 781 AGCTCTCTGGGACACTCTGTACTGAGCAGACTGGCAGACCGGCTCCATCGCTGTC 840
Db 829 AGCTCTCTGGGACACTCTGTACTGAGCAGACTGGCAGACCGGCTCCATCGCTGTC 888
QY 841 AACAGCGCACTGGGGGAGAGGAGAGATCTCTGAGTGCCTCTACTCAACCATGAGC 900
Db 889 AACAGCGCACTGGGGGAGAGGAGAGATCTCTGAGTGCCTCTACTCAACCATGAGC 948
QY 901 ATCCAGGTGCTGAGCCAGGAGCGGAGCCTTTTCCACACTCGCTGTGAGGAGCAAT 960
Db 949 ATCCAGGTGCTGAGCCAGGAGCGGAGCCTTTTCCACACTCGCTGTGAGGAGCAAT 1008
QY 961 GCGGCTGCTCCACCTGTGCTGTCTGTCCTCCAGCGAGCCTTCTACATGCGCTGTC 1020
Db 1009 GCGGCTGCTCCACCTGTGCTGTCTGTCCTCCAGCGAGCCTTCTACATGCGCTGTC 1068
QY 1021 CCCACGGGTGTCAGCTGCAGACAAACGCGAGGAGCTGTAGGAGGAGCGGAGAGTG 1080
Db 1069 CCCACGGGTGTCAGCTGCAGACAAACGCGAGGAGCTGTAGGAGGAGCGGAGAGTG 1128
QY 1081 CTGCTGTGCGCGCGGAGCGGACCTACGAGGATCTCGCTGGGACAGCGCGGACTTCAAC 1140
Db 1129 CTGCTGTGCGCGCGGAGCGGACCTACGAGGATCTCGCTGGGACAGCGCGGACTTCAAC 1188
QY 1141 GACATCGTGTGAGGTGGAGCAGATCCGACCGCCATGCGCTGAGCTAGACCGGCTA 1200
Db 1189 GACATCGTGTGAGGTGGAGCAGATCCGACCGCCATGCGCTGAGCTAGACCGGCTA 1248
QY 1201 GAGGCTATGTCTACTGACAGATGACGAGGTGGGGCCATCCGACGGGCGCTACTGAGC 1260
Db 1249 GAGGCTATGTCTACTGACAGATGACGAGGTGGGGCCATCCGACGGGCGCTACTGAGC 1308
QY 1261 GGGTCTGGGGCGCAGCGCTGGTCAACCGAGATCAACGACCCCGATGGCATCGCGTC 1320

Db 1309 GGGTCTGGGGCGCAGACGCTGGTCAACACCGAGATCAACGACCCCGATGGCATCCGGTC 1368
QY 1321 GACTGGGTGGCCGAAACCTCTACTGGACCGACACCGGCAACCGCATCGAGGTGAGC 1380
Db 1369 GACTGGGTGGCCGAAACCTCTACTGGACCGACACCGGCAACCGCATCGAGGTGAGC 1428
QY 1381 CGCTCAACCGGACCTCCCGCAAGATCTGTGTGCGAGGACCTGGAGAGCCCGGAGCC 1440
Db 1429 CGCTCAACCGGACCTCCCGCAAGATCTGTGTGCGAGGACCTGGAGAGCCCGGAGCC 1488
QY 1441 ATCGCACTGCAACCGCTGATGGGCTCATGTACTGGACAGACTGGGAGAGAACCTTAA 1500
Db 1489 ATCGCACTGCAACCGCTGATGGGCTCATGTACTGGACAGACTGGGAGAGAACCTTAA 1548
QY 1501 ATCGAGTGTGCAACTTTGGATGGGCGAGGCGGGTGTGTGGTCAATTCCTCCCTCGGG 1560
Db 1549 ATCGAGTGTGCAACTTTGGATGGGCGAGGAGCGGCTGTGTGGTCAATTCCTCCCTCGGG 1608
QY 1561 TGGCCCAACCGGCTGGCCCTGGACCTGACAGAGGGGAGCTCTACTGGGAGAGCCCAAG 1620
Db 1609 TGGCCCAACCGGCTGGCCCTGGACCTGACAGAGGGGAGCTCTACTGGGAGAGCCCAAG 1668
QY 1621 ACAGCAAGATCGAGGTGATCAATGTTGATGGGACGAGGCGGACCTCTCTGGAGAC 1680
Db 1669 ACAGCAAGATCGAGGTGATCAATGTTGATGGGACGAGGCGGACCTCTCTGGAGAC 1728
QY 1681 AAGCTCCGCAATTTTGGGTTTCACTGTCTGGGGGACTTCTACTTGGAGTGAATCTG 1740
Db 1729 AAGCTCCGCAATTTTGGGTTTCACTGTCTGGGGGACTTCTACTTGGAGTGAATCTG 1788
QY 1741 CAGCGCGCAGCATCGAGCGGCTGCAAGGTGCAAGGTGCAAGGTGCAAGGTGCAAGGTGCA 1800
Db 1789 CAGCGCGCAGCATCGAGCGGCTGCAAGGTGCAAGGTGCAAGGTGCAAGGTGCAAGGTGCA 1848
QY 1801 CAGCTGCGGACCTGATGGGCTCAAGGTGCAAGGTGCAAGGTGCAAGGTGCAAGGTGCA 1860
Db 1849 CAGCTGCGGACCTGATGGGCTCAAGGTGCAAGGTGCAAGGTGCAAGGTGCAAGGTGCA 1908
QY 1861 CCGTGTGGGACAGAAACGGGGGTGAGCCACTGTGTCTTTCACACCCAGCAAC 1920
Db 1909 CCGTGTGGGACAGAAACGGGGGTGAGCCACTGTGTCTTTCACACCCAGCAAC 1968
QY 1921 CCGTGTGGTCCCGCATCGGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1980
Db 1969 CCGTGTGGTCCCGCATCGGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 2028
QY 1981 GAGGCTCTTGTGTCTTTCACAGCAGAGCGGCTTCCACAGGATCTCCCTCGAGACCAAT 2040
Db 2029 GAGGCTCTTGTGTCTTTCACAGCAGAGCGGCTTCCACAGGATCTCCCTCGAGACCAAT 2088
QY 2041 AACAAACAGCTGGGCTCCGCTCACGGGGTCAAGGAGGCTCAGGCTTGGACTTTGAT 2100
Db 2089 AACAAACAGCTGGGCTCCGCTCACGGGGTCAAGGAGGCTCAGGCTTGGACTTTGAT 2148
QY 2101 GTGTCCAAACACACATCTACTGGACAGAGCTGAGCTGAGGACCATCAGCGCGGCTTC 2160
Db 2149 GTGTCCAAACACACATCTACTGGACAGAGCTGAGCTGAGGACCATCAGCGCGGCTTC 2208
QY 2161 ATGAAACGGGAGCTGGTGGAGCAGCTGTGGAGTTTGGCTTGAATCCCGAGGCGATG 2220
Db 2209 ATGAAACGGGAGCTGGTGGAGCAGCTGTGGAGTTTGGCTTGAATCCCGAGGCGATG 2268
QY 2221 GCGCTTCACTGGATGGGCAAGACCTCTACTGGGCGGACCTGGGACCAACAGAAATCGAA 2280
Db 2269 GCGCTTCACTGGATGGGCAAGACCTCTACTGGGCGGACCTGGGACCAACAGAAATCGAA 2328
QY 2281 GTGCGCGGCTGGACCGGCAAGTTCCGCAAGTCTCTGTGTGGAGGACTTGGACAAACCG 2340
Db 2329 GTGCGCGGCTGGACCGGCAAGTTCCGCAAGTCTCTGTGTGGAGGACTTGGACAAACCG 2388
QY 2341 AGTGTGCTGGCCCTGGATCCCAACAGGGGCTACATCTACTGGACCGAGTGGGGGCGAG 2400
Db 2389 AGTGTGCTGGCCCTGGATCCCAACAGGGGCTACATCTACTGGACCGAGTGGGGGCGAG 2448

QY 2401 CCGAGGATCGTGGGCTTTCATGACCGGACCAACTGCATGACGCTGGTGCAAGGTG 2460
DB 2449 CCGAGGATCGTGGGCTTTCATGACCGGACCAACTGCATGACGCTGGTGCAAGGTG 2508
QY 2461 GGGCGGGCAACGACCTCACATTTGACTACGCTGACAGCGCTCTACTGACGACCTG 2520
DB 2509 GGGCGGGCAACGACCTCACATTTGACTACGCTGACAGCGCTCTACTGACGACCTG 2568
QY 2521 GACACCAATGATGACGCTGCTCAACATGCTGGGTGAGGAGCGGTGCTGATGGCGAC 2580
DB 2569 GACACCAATGATGACGCTGCTCAACATGCTGGGTGAGGAGCGGTGCTGATGGCGAC 2628
QY 2581 GATCTCCGCGCACCGCTTGGTCTGACGAGTACAGCGATTATATCTACTGACAGACTGG 2640
DB 2629 GATCTCCGCGCACCGCTTGGTCTGACGAGTACAGCGATTATATCTACTGACAGACTGG 2688
QY 2641 AATCTGCACAGATTGAGCGGGCGGACAAAGCTAGCGCGCGGAAACCGACCTCATCCAG 2700
DB 2689 AATCTGCACAGATTGAGCGGGCGGACAAAGCTAGCGCGCGGAAACCGACCTCATCCAG 2748
QY 2701 GGGCACCTGGACTTGTGATGAGATCTCTGGTGTTCACATCTCCCGCAGGATGGCTC 2760
DB 2749 GGGCACCTGGACTTGTGATGAGATCTCTGGTGTTCACATCTCCCGCAGGATGGCTC 2808
QY 2761 AATGACTGTATGCAACAAACGGGCGAGTGTGGGCGAGCTGTGCTTGCCTATCCCGGGCGC 2820
DB 2809 AATGACTGTATGCAACAAACGGGCGAGTGTGGGCGAGCTGTGCTTGCCTATCCCGGGCGC 2868
QY 2821 CACCGCTGCGGTGCGCTTCACTACATCACTGTGACCCGACGCGCAACTGACGCGCG 2880
DB 2869 CACCGCTGCGGTGCGCTTCACTACATCACTGTGACCCGACGCGCAACTGACGCGCG 2928
QY 2881 CCCACCACTTCTGTGTTGAGCCAGCAAACTGCCATCAGTCCGATGATCCCGGACGAC 2940
DB 2929 CCCACCACTTCTGTGTTGAGCCAGCAAACTGCCATCAGTCCGATGATCCCGGACGAC 2988
QY 2941 CAGCACAGCCCGGATCTCATCTGCTCCCTGCATGGACTGAGGAAAGTCAAGACCATCGAC 3000
DB 2989 CAGCACAGCCCGGATCTCATCTGCTCCCTGCATGGACTGAGGAAAGTCAAGACCATCGAC 3048
QY 3001 TATGACCACTGGAAGTTTCACTACCTGAGTGGGATGGGCGGCAAGACATCAAGCGAGCC 3060
DB 3049 TATGACCACTGGAAGTTTCACTACCTGAGTGGGATGGGCGGCAAGACATCAAGCGAGCC 3108
QY 3061 AAGGACGAGCGGACCCAGCCCTTTGTTTGGACCTCTGAGCGCAAGGCAAAACCCAGAC 3120
DB 3109 AAGGACGAGCGGACCCAGCCCTTTGTTTGGACCTCTGAGCGCAAGGCAAAACCCAGAC 3168
QY 3121 AGGACGCGCCACGACCTCAGCATCGACATCTACAGCGGACACTGTTCTGAGCGTGGCGAG 3180
DB 3169 AGGACGCGCCACGACCTCAGCATCGACATCTACAGCGGACACTGTTCTGAGCGTGGCGAG 3228
QY 3181 GCCACCAATACCATCAACGTTCCACAGGCTGAGCGGGGAAGCCATGGGGGTGGTCTGCGT 3240
DB 3229 GCCACCAATACCATCAACGTTCCACAGGCTGAGCGGGGAAGCCATGGGGGTGGTCTGCGT 3288
QY 3241 GGGGACCGCGACAGCCCGAGGCGCATCGTCTGACAGCGGAGCGAGGCTACTGTTACTTC 3300
DB 3289 GGGGACCGCGACAGCCCGAGGCGCATCGTCTGACAGCGGAGCGAGGCTACTGTTACTTC 3348
QY 3301 ACCAACATGACGAGCCCGGACGCAAGATCGAAACGCGCAGCCCTGGACCGGACCGAGCGC 3360
DB 3349 ACCAACATGACGAGCCCGGACGCAAGATCGAAACGCGCAGCCCTGGACCGGACCGAGCGC 3408
QY 3361 GAGGTCTCTTTCACCAACCGGCTCATCCGCGCTGTGGCGCTGGTGGTGGACAAACACTG 3420
DB 3409 GAGGTCTCTTTCACCAACCGGCTCATCCGCGCTGTGGCGCTGGTGGTGGACAAACACTG 3468
QY 3421 GGCAGCTGTTCTGGTGGACCGCGACCTGAGCGCATTTGAGAGCTGTGACCTGTCAAGG 3480
DB 3469 GGCAGCTGTTCTGGTGGACCGCGACCTGAGCGCATTTGAGAGCTGTGACCTGTCAAGG 3528

QY 3481 GCCAACCGCCTGACCCCTGGAGGACGCCAACATCGTGCAGCCTCTGGGCTGACCATCCTT 3540
DB 3529 GCCAACCGCCTGACCCCTGGAGGACGCCAACATCGTGCAGCCTCTGGGCTGACCATCCTT 3588
QY 3541 GGCAGGATCTCTACTGATCGACCGGACGACGAGATGATCGAGCGTGTGGAGAGACC 3600
DB 3589 GGCAGGATCTCTACTGATCGACCGGACGACGAGATGATCGAGCGTGTGGAGAGACC 3648
QY 3601 ACCGGGACAAAGCGGACTCGCATCCAGGCGGTGTGGCCACCTCCTCAGTGGCATCCATGCA 3660
DB 3649 ACCGGGACAAAGCGGACTCGCATCCAGGCGGTGTGGCCACCTCCTCAGTGGCATCCATGCA 3708
QY 3661 GTGAGGAGAGTCACTCTGGAGGAGTCTTCAAGCCCAACCATGTGCCGCTGACAAATGGTGGC 3720
DB 3709 GTGAGGAGAGTCACTCTGGAGGAGTCTTCAAGCCCAACCATGTGCCGCTGACAAATGGTGGC 3768
QY 3721 TGCTCCACATCTGTATTTGCCAAGGGTGTGGGACACCAACGGTGTCTATGCCAGTCCAC 3780
DB 3769 TGCTCCACATCTGTATTTGCCAAGGGTGTGGGACACCAACGGTGTCTATGCCAGTCCAC 3828
QY 3781 CTGCTGTCTTGCAGAACTCTGACCTGTGGAGAGCGGCCACCTGCTCCCGGACCGAC 3840
DB 3829 CTGCTGTCTTGCAGAACTCTGACCTGTGGAGAGCGGCCACCTGCTCCCGGACCGAC 3888
QY 3841 TTTGATGTGTCACAGGGGAGATCGACTATATCCCGGGGCTTGGCGCTGTGACGGCTTT 3900
DB 3889 TTTGATGTGTCACAGGGGAGATCGACTATATCCCGGGGCTTGGCGCTGTGACGGCTTT 3948
QY 3901 CCCGAGTGCAGTACAGAGCGAGAGGGCTGCCCGGTGTGCTTCCCGCGCCAGTTC 3960
DB 3949 CCCGAGTGCAGTACAGAGCGAGAGGGCTGCCCGGTGTGCTTCCCGCGCCAGTTC 4008
QY 3961 CCTTGGCGGGGGTCACTGTGTGACCTGCGCTGCGCTGCGAGCGGAGGAGACTGT 4020
DB 4009 CCTTGGCGGGGGTCACTGTGTGACCTGCGCTGCGCTGCGAGCGGAGGAGACTGT 4068
QY 4021 CAGGACCGCTCAGACGAGGGGAGTGTGACGCTCTGCTTCCGCGGCAACAGTTCGGGTGT 4080
DB 4069 CAGGACCGCTCAGACGAGGGGAGTGTGACGCTCTGCTTCCGCGGCAACAGTTCGGGTGT 4128
QY 4081 GCGAGCGGCGAGTGTGCTCTCATCAACAGCAGTGTGAGTCTTCTCCCGGACTGTATCGAC 4140
DB 4129 GCGAGCGGCGAGTGTGCTCTCATCAACAGCAGTGTGAGTCTTCTCCCGGACTGTATCGAC 4188
QY 4141 GGCTCCGACGAGCTCATGTGAAATCAACAAAGCGGCTCAGACGACGAGCGGCGCCAC 4200
DB 4189 GGCTCCGACGAGCTCATGTGAAATCAACAAAGCGGCTCAGACGACGAGCGGCGCCAC 4248
QY 4201 AGCAGTGCCTATCGGGCGGCTCATTTGGCATCATCTCTCTCTTCTGTCATGGGTGGTGC 4260
DB 4249 AGCAGTGCCTATCGGGCGGCTCATTTGGCATCATCTCTCTCTTCTGTCATGGGTGGTGC 4308
QY 4261 TATTTTGTGTGCGAGCGGTGTGTCAGGGCTATGGGGGGGCAACGGGCGCTTCCCG 4320
DB 4309 TATTTTGTGTGCGAGCGGTGTGTCAGGGCTATGGGGGGGCAACGGGCGCTTCCCG 4368
QY 4321 CACGAGTATGTACGCGGAGCCCGCGCTGCTCAATTTTCATAGCCCGGGGGGTTC 4380
DB 4369 CACGAGTATGTACGCGGAGCCCGCGCTGCTCAATTTTCATAGCCCGGGGGGTTC 4428
QY 4381 CAGCATGGCGCTTTCACAGGCGATCGCATGGGAAAGTCCATGATGAGCTCCGTGAGCGCTG 4440
DB 4429 CAGCATGGCGCTTTCACAGGCGATCGCATGGGAAAGTCCATGATGAGCTCCGTGAGCGCTG 4488
QY 4441 ATGGGGGGCGGGGGGGTGGCTCTGTATGAAACGCGAAACAGTCAAGGGGCTCTGTC 4500
DB 4489 ATGGGGGGCGGGGGGGTGGCTCTGTATGAAACGCGAAACAGTCAAGGGGCTCTGTC 4548
QY 4501 AGCAGCTGTCTCAGACGAGAGGCGCACTGTATCCCGCGGATCTCTGAAACCGCGCGCTCC 4560
DB 4549 AGCAGCTGTCTCAGACGAGAGGCGCACTGTATCCCGCGGATCTCTGAAACCGCGCGCTCC 4608
QY 4561 CCGGCGCACGGACCCCTCCCTGTACAAACATGACATGTTTCTACTCTTCAAAACATTCGGGCC 4620

4609	Db	CCGCGCCACGGACCCCTCCCTGTGTAACAATGGACATGTTCTACTCTTCAAACATCTTCGGCC	4668
4621	Qy	ACTGCGAGACCGTACAGCGCCCTACATCATTTGAGGAAATGCGGCCCGCCGACGACGCCCTGC	4680
4669	Db	ACTGCGAGACCGTACAGCGCCCTACATCATTTGAGGAAATGCGGCCCGCCGACGACGCCCTGC	4728
4681	Qy	AGCACCGACGTGTGTGACAGCGACTACAGCGCCAGCGCGTGTGGAAGGCCACGAAGTACTAC	4740
4729	Db	AGCACCGACGTGTGTGACAGCGACTACAGCGCCAGCGCGTGTGGAAGGCCACGAAGTACTAC	4788
4741	Qy	CTGGATTTGAACTCGGACTCAGACCCCTATCCACCCCAACCGCCACCGCCACGCCAGTAC	4800
4789	Db	CTGGATTTGAACTCGGACTCAGACCCCTATCCACCCCAACCGCCACCGCCAGTAC	4848
4801	Qy	CTGTGCGCGGAGACAGCTGCGCGCCCTCGCGCGCCACCGAGAGGAGTACTTCCATCTC	4860
4849	Db	CTGTGCGCGGAGACAGCTGCGCGCCCTCGCGCGCCACCGAGAGGAGTACTTCCATCTC	4908
4861	Qy	TTCCCGCCCCCTCGTCCCGCTGCA CGGACTCATCTGAACTCTGCGCGCGGCCACTCTGGC	4920
4909	Db	TTCCCGCCCCCTCGTCCCGCTGCA CGGACTCATCTGAACTCTGCGCGCGGCCACTCTGGC	4968
4921	Qy	TTCTCTGTGCGCCCTGTAATAAGTTTTTAAATATGAACAAAGAAAAATATATTTATGAT	4980
4969	Db	TTCTCTGTGCGCCCTGTAATAAGTTTTTAAATATGAACAAAGAAAAATATATTTATGAT	5028
4981	Qy	TTAAAAAATAAATAAATTTGGATTTTTAAAAACAATGAGAAATGTGAACCTGTGATGGGGTG	5040
5029	Db	TTAAAAAATAAATAAATTTGGATTTTTAAAAACAATGAGAAATGTGAACCTGTGATGGGGTG	5088
5041	Qy	GGCAGGGCTGGAGAACTTTTGTA	5063
5089	Db	GGCAGGGCTGGAGAACTTTTGTA	5111

Search completed: February 19, 2005, 06:21:24
Job time : 2609 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 17, 2005, 01:15:53 ; Search time 201 seconds
(without alignments)
3107.552 Million cell updates/sec

Title: US-09-931-375A-2

Perfect score: 8740

Sequence: 1 MEAAPGPPWPLLLLLLLLLL.....TERSYFHLPPPPSPCTDSS 1615

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq_16Dec04:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8740	100.0	1615	5 AAE21730	AAE21730 Human bon
2	8740	100.0	1615	6 ABR41131	ABr41131 Human LRP
3	8740	100.0	1615	7 ADB98798	ADb98798 Human Zma
4	8740	100.0	1615	8 ABO84659	ABo84659 Human can
5	8736	100.0	1615	2 AAWS3309	AAw83309 LRPS prot
6	8736	100.0	1615	5 AAE21740	AAe21740 Human BSM
7	8736	100.0	1615	8 ADI27181	ADi27181 Human LRP
8	8736	100.0	1615	8 ADQ20524	ADq20524 Human sof
9	8736	100.0	1615	8 ADR17560	ADr17560 Human hig
10	8736	100.0	1615	8 ADR48211	ADr48211 Human hig
11	8736	100.0	1615	8 ADR73482	ADr73482 Human low
12	8735	99.9	1615	5 AAE21741	AAe21741 Human BSM
13	8727	99.9	1615	8 ADR17561	ADr17561 Human hig
14	8727	99.9	1615	8 ADR48212	ADr48212 Human hig
15	8724	99.8	1627	7 ABR85665	ABr85665 Human pro
16	8724	99.8	1627	8 ABO84660	ABo84660 Human can
17	8715	99.7	1615	4 AAG68169	AGg68169 Human Zma
18	8715	99.7	1615	6 ABR41093	ABr41093 Human wil
19	8715	99.7	1615	7 ADB98058	ADb98058 Human LRP
20	8715	99.7	1615	7 ADE82427	ADe82427 Human Zma
21	8715	99.7	1615	8 ADR16921	ADr16921 Human hig
22	8715	99.7	1615	8 ADR47572	ADr47572 Human hig
23	8706	99.6	1615	4 AAG68170	AGg68170 Human HBM
24	8706	99.6	1615	5 AAU80880	AAu80880 Human hig
25	8706	99.6	1615	6 ABR41094	ABr41094 Human LRP

26	8706	99.6	1615	7 ADB98059	ADb98059 LRPS mute
27	8706	99.6	1615	7 ADE82428	ADe82428 Human HBM
28	8706	99.6	1615	8 ADR16922	ADr16922 Human hig
29	8706	99.6	1615	8 ADR47573	ADr47573 Human hig
30	8705	99.6	1611	8 ADI27180	ADi27180 Human LRP
31	8705	99.6	1665	6 ABR41133	ABr41133 Human LRP
32	8705	99.6	1665	7 ADB98800	ADb98800 Human Zma
33	8702	99.6	1615	5 AAU80879	AAu80879 Human Zma
34	8607	98.5	1591	2 AAWS3308	AAw83308 Mature LRP
35	8574	98.1	1639	2 AAWS3311	AAw83311 LRPS isof
36	8259.5	94.5	1614	2 AAWS3312	AAw83312 Mouse LRP
37	8259.5	94.5	1614	5 ABB07255	ABb07255 Mouse LRP
38	8259.5	94.5	1614	8 ADI27179	ADi27179 Mouse LRP
39	8254.5	94.4	1614	8 ADI27193	ADi27193 Mouse LRP
40	8254.5	94.4	1614	8 ADI27174	ADi27174 Mouse LRP
41	7972.5	91.2	1564	6 ABR41132	ABr41132 Mouse LRP
42	7972.5	91.2	1564	7 ADB98799	ADb98799 Mouse Zma
43	7882	90.2	1451	2 AAWS3310	AAw83310 LRPS prot
44	6868	78.6	1325	8 ABO84658	ABo84658 Mouse can
45	6153.5	70.4	1613	6 ABR41134	ABr41134 Human LRP

ALIGNMENTS

RESULT 1
AAE21730
ID AAE21730 standard; protein; 1615 AA.
XX AC AAE21730;
XX 16-JUL-2002 (first entry)
DT
XX
DE Human bone strength and mineralisation regulatory protein (BSMR).
KW Human; bone strength and mineralisation regulatory protein; BSMR;
KW bone strength; mineralisation; ophthalmological; antidiabetic;
KW bone density regulating transmembrane receptor; prosthetic device;
KW surgical implant; diabetic retinopathy; hypertensive retinopathy;
KW therapy; osteoporosis; prematurity; ocular vessel; eye disorder;
KW osteopathic.
XX
OS Homo sapiens.
XX
PN WO200216553-A2.
XX
PD 28-FEB-2002.
XX
PF 17-AUG-2001; 2001WO-US041788.
XX
PR 18-AUG-2000; 2000US-0226119P.
PR 22-SEP-2000; 2000US-0234337P.
PR 13-JUL-2001; 2001US-0304851P.
XX
PA (AVET) AVENTIS PHARMA SA.
PA (HARD) HARVARD COLLEGE.
PA (UYCA-) UNIV CASE WESTERN RESERVE.
XX
PI Warman ML, Gong Y, Olsen BR, Rawadi G, Roman-Roman S;
XX WPI; 2002-329694/36.
XX N-PSDB; AAD34322.
XX
PT Polynucleotide encoding bone strength and mineralization regulatory
XX protein useful for diagnosis or therapy of osteoporosis.
XX Disclosure; Fig 3; 124pp; English.
XX
CC The invention relates to bone strength and mineralisation regulatory
XX protein (BSMR) and its corresponding nucleic acid sequence. BSMR DNA is
XX useful for the diagnosis or therapy of osteoporosis and for regulating
XX (increasing) bone strength and mineralisation in a human subject by
XX activating a bone density regulating transmembrane receptor (BSMR)

protein). An expression vector comprising a promoter that is operably linked to BSMR DNA is useful for modulating bone density and for enhancing bone strength and mineralisation in a mammal cell. Composition comprising a BSMR effector is useful for treating osteoporosis and is useful particularly as a coating for prosthetic devices and surgical implants. BSMR is useful for screening lead pharmaceutical agents as BSMR effectors, which may be used to treat a range of eye disorders such as diabetic retinopathy, hypertensive retinopathy and retinopathy of prematurity, in which normal vascular growth and integrity of ocular vessels is disrupted. The present sequence is human BSMR protein

XX Sequence 1615 AA;

Query Match
Best Local Similarity 100.0%; Score 8740; DB 5; Length 1615;
Matches 1615; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEAAPGPPWPLLLLLLLLLLALCGCPAPAAASPLLLFANRRDRLVLDAGGVKLESTIVVS 60
DB 1 MEAAPGPPWPLLLLLLLLLLALCGCPAPAAASPLLLFANRRDRLVLDAGGVKLESTIVVS 60

QY 61 GLEDAADVDFQSKGAVYTDVSEBAIKQTYLNQTAAGVQNVVLSGLVSPDGLACDWVGK 120
DB 61 GLEDAADVDFQSKGAVYTDVSEBAIKQTYLNQTAAGVQNVVLSGLVSPDGLACDWVGK 120

QY 121 KLYTDSNTNRELVANLNGTSRKVLFWQDLOPRAIALDPAGYNYWTDGTPRIERAG 180
DB 121 KLYTDSNTNRELVANLNGTSRKVLFWQDLOPRAIALDPAGYNYWTDGTPRIERAG 180

QY 181 MDGSTRKIIIVSDIYWPNGLTIDLEEQKLYWADAKLSFIHRANDGSPRKQVVEGSLTHP 240
DB 181 MDGSTRKIIIVSDIYWPNGLTIDLEEQKLYWADAKLSFIHRANDGSPRKQVVEGSLTHP 240

QY 241 FALTLSGDTLWTDQWTSIHACNKRGTGKKEILSALYSPMDIQVLQSERQPPFHTTCE 300
DB 241 FALTLSGDTLWTDQWTSIHACNKRGTGKKEILSALYSPMDIQVLQSERQPPFHTTCE 300

QY 301 EDNGGCSHLCLLSPSEPYTCACPTGVQLONGRTCKAGAEVLLIARLTDLRLSLDTP 360
DB 301 EDNGGCSHLCLLSPSEPYTCACPTGVQLONGRTCKAGAEVLLIARLTDLRLSLDTP 360

QY 361 DFTDVLQVDDIRHAIADYDPLEGYVYTTDDEVRAIRRAYLDGSGAQLVNTINDDPG 420
DB 361 DFTDVLQVDDIRHAIADYDPLEGYVYTTDDEVRAIRRAYLDGSGAQLVNTINDDPG 420

QY 421 IAVDWARNLYWTDGTDRIEVTRLNGTSRKILVSEDLDEPRALHPVWGLMTWTDGCE 480
DB 421 IAVDWARNLYWTDGTDRIEVTRLNGTSRKILVSEDLDEPRALHPVWGLMTWTDGCE 480

QY 481 NPKIECANLDQERRVLYNVAISLGNPGLALDLOEGKLYWDAKTDKIEVINVDGTKRRTL 540
DB 481 NPKIECANLDQERRVLYNVAISLGNPGLALDLOEGKLYWDAKTDKIEVINVDGTKRRTL 540

QY 541 LEDKLPHPITGLDGFYIYTDWQRRSIRVHKVKSADVIIDQLPDLMLGLKAVNVAKV 600
DB 541 LEDKLPHPITGLDGFYIYTDWQRRSIRVHKVKSADVIIDQLPDLMLGLKAVNVAKV 600

QY 601 GTNPCADRNGCSHLCPFTPHATRCGCPIGLELLSDMKTCIVPEAFVFTSRAAIHRLS 660
DB 601 GTNPCADRNGCSHLCPFTPHATRCGCPIGLELLSDMKTCIVPEAFVFTSRAAIHRLS 660

QY 661 ETNNNDVAIPLTGVKEASALDFDVSNHIIYWDVSLKTIISRAFWNGSSVHVVEFGLDYP 720
DB 661 ETNNNDVAIPLTGVKEASALDFDVSNHIIYWDVSLKTIISRAFWNGSSVHVVEFGLDYP 720

QY 721 EGMADVMMGKLYWADTGTNRIEVARLDQGRQVLVWRDLNPRSLALDPTKGYIYWTWE 780
DB 721 EGMADVMMGKLYWADTGTNRIEVARLDQGRQVLVWRDLNPRSLALDPTKGYIYWTWE 780

QY 781 GGPRIVRAFMDFGNCMTLVKVGANDLTIDYADQRLYWTDLTNMIESNMILQSERV 840
DB 781 GGPRIVRAFMDFGNCMTLVKVGANDLTIDYADQRLYWTDLTNMIESNMILQSERV 840

QY 841 IADDLPHFPGLTOYSDIYIYTDWNLHSLERADKTSGRNRTLIQGHLDVMDILVPHSSRQ 900
DB 841 IADDLPHFPGLTOYSDIYIYTDWNLHSLERADKTSGRNRTLIQGHLDVMDILVPHSSRQ 900

QY 901 DGLNDGMNNGCCQQLCIAIPGHRCCGASHYTLDPSSRNCSPPTTLLFSQKSAISRWI 960
DB 901 DGLNDGMNNGCCQQLCIAIPGHRCCGASHYTLDPSSRNCSPPTTLLFSQKSAISRWI 960

QY 961 PDDQHSPLIILPLHGLRNVAKAIDYDPLDKFYIYWDGRQNIKRAKDDGTQPPVLTSLSQG 1020
DB 961 PDDQHSPLIILPLHGLRNVAKAIDYDPLDKFYIYWDGRQNIKRAKDDGTQPPVLTSLSQG 1020

QY 1021 NPDQPHDLSIDIYSRTLFWTCEATNTINVRHLSGEAMGVVLRGDRDKPRAIVNAERG 1080
DB 1021 NPDQPHDLSIDIYSRTLFWTCEATNTINVRHLSGEAMGVVLRGDRDKPRAIVNAERG 1080

QY 1081 LYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLFWVDADLKRIESCD 1140
DB 1081 LYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLFWVDADLKRIESCD 1140

QY 1141 LSGANRLTLEDANIVQPLGLTILGKHLVYIDRQOQMIERVEKTTGDKRTRIQGRVAHLTG 1200
DB 1141 LSGANRLTLEDANIVQPLGLTILGKHLVYIDRQOQMIERVEKTTGDKRTRIQGRVAHLTG 1200

QY 1201 IHAVEEVSLEEFSAHPCARDNGGCSHICIAKGDGTPRCSCPVLHVLVONLLITCGEPPTCS 1260
DB 1201 IHAVEEVSLEEFSAHPCARDNGGCSHICIAKGDGTPRCSCPVLHVLVONLLITCGEPPTCS 1260

QY 1261 PDQFACATGEIDCIPGAWRCDFPECCDDQSDDEGCPVCSAAQFFCARGQCVDLRLCDGE 1320
DB 1261 PDQFACATGEIDCIPGAWRCDFPECCDDQSDDEGCPVCSAAQFFCARGQCVDLRLCDGE 1320

QY 1321 ADCQDRSEADCAICLPNCFRASCQCVLIKQCCDFFPCIDGSDLMCEITKPPSDDS 1380
DB 1321 ADCQDRSEADCAICLPNCFRASCQCVLIKQCCDFFPCIDGSDLMCEITKPPSDDS 1380

QY 1381 PAHSSAIGPVIGIITLSLFWMGVYFVQVVCQRYAGANGPPHVEYVSGTHPVLNFTAP 1440
DB 1381 PAHSSAIGPVIGIITLSLFWMGVYFVQVVCQRYAGANGPPHVEYVSGTHPVLNFTAP 1440

QY 1441 GSGHGGPFTGIACGKSNMSSVSLMGGRGVPLVDRNHVTGASSSSSSSTKATLYPPIINP 1500
DB 1441 GSGHGGPFTGIACGKSNMSSVSLMGGRGVPLVDRNHVTGASSSSSSSTKATLYPPIINP 1500

QY 1501 PPSPATDPSLYNMDFYSSNIPATARYPIIIRMGMAPPTTTCSTDVCDSDYSASRWKAS 1560
DB 1501 PPSPATDPSLYNMDFYSSNIPATARYPIIIRMGMAPPTTTCSTDVCDSDYSASRWKAS 1560

QY 1561 KYLDLNSDSDPYPPPTPHSQVLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615
DB 1561 KYLDLNSDSDPYPPPTPHSQVLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615

RESULT 2

ABR41131
ID ABR41131 standard; protein; 1615 AA.

AC ABR41131;

XX 02-JUN-2003 (first entry)

XX Human LRP5 protein.

XX Human; high bone mass; HEM; LRP5; LRP6; transgenic; bone mass modulation;
KW gene therapy; bone density modulation; bone strength; trabecular number;
KW bone size; bone tissue connectivity; bone disease; osteoporosis;
KW osteomalacia; rickets; Paget's disease; neoplasm of the bone.

OS Homo sapiens.

XX WO200292764-A2.

XX 21-NOV-2002.

XX PF 13-MAY-2002; 2002WO-US014876.
XX PR 11-MAY-2001; 2001US-0290071P.
XX PR 17-MAY-2001; 2001US-0291311P.
XX PR 01-FEB-2002; 2002US-0353058P.
XX PR 04-MAR-2002; 2002US-0361293P.
XX PA (GENO-) GENOME THERAPEUTICS CORP.
XX PA (AMHP) WYETH.
XX PI Babij P, Bex FJ, Yaworsky PJ, Bodine PV;
XX WPI; 2003-129278/12.
XX New transgenic animals (e.g. mice), useful as models for studying bone
PT density modulation, developing drugs for treating or preventing bone
PT diseases (e.g. osteoporosis), or diagnosing diseases characterized by
PT reduced bone density.
XX Disclosure; Fig 26; 603pp; English.
XX
CC The invention relates to novel transgenic animals expressing the high
CC bone mass (HBM) gene, expressing the corresponding wild type HBM gene,
CC comprising an alteration of the gene encoding LRP5 or LRP6, or expressing
CC an LRP5 that is modulated by an altered gene control sequence introduced
CC by homologous or non-homologous recombination. The transgenic animals are
CC for the study of bone density modulation or bone mass modulation. The
CC invention has osteopathic and cytostatic activity. The polynucleotides of
CC the invention may have a use in gene therapy. The transgenic animals and
CC nucleic acids are for the study of bone density modulation, where the
CC bone mass is modulated relative to non-transgenic animals of the same
CC species in more than one parameter selected from bone density, bone
CC strength, trabecular number, bone size, or bone tissue connectivity. The
CC transgenic animals, nucleic acids and methods are useful for identifying
CC molecules involved in bone development, and for developing pharmaceutical
CC compositions, which may be employed for treating or preventing bone
CC diseases, e.g. osteoporosis, osteomalacia, rickets, Paget's disease, or
CC neoplasms of the bone. The transgenic animals and nucleic acids are also
CC useful in methods for diagnosing diseases involved in bone development, is
CC or characterised by reduced bone density or mass. The present sequence, is
XX used in the exemplification of the invention
SQ Sequence 1615 AA;
Query Match 100.0%; Score 8740; DB 6; Length 1615;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1615; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MEAAPPGPWPLLILLILLALCGCPAPAAASPLLFPANRRDRLVLDAGGVKLESTIVS 60
DB 1 MEAAPPGPWPLLILLILLALCGCPAPAAASPLLFPANRRDRLVLDAGGVKLESTIVS 60
QY 61 GLEDAAAVDFQSKGAVTVDVSEAIKQTYLNTQTGAQVNVVIGSLVSPDGLACDWGK 120
DB 61 GLEDAAAVDFQSKGAVTVDVSEAIKQTYLNTQTGAQVNVVIGSLVSPDGLACDWGK 120
QY 121 KLYWTDSETNRIEVANLNGTSRKVLFWQDLQOPRAIALDPAGHYMYTWDWGETPRIERAG 180
DB 121 KLYWTDSETNRIEVANLNGTSRKVLFWQDLQOPRAIALDPAGHYMYTWDWGETPRIERAG 180
QY 181 MDGSTRKIIVDSIYWPNGLIIDLEOKLYWADAKLSFIHRANLDGSPROKRVGSLTHP 240
DB 181 MDGSTRKIIVDSIYWPNGLIIDLEOKLYWADAKLSFIHRANLDGSPROKRVGSLTHP 240
QY 241 FALTSLGDTLWTQWTRSIHACNKRGTGGRKEILSALYSPMDIQVLSQERQPFHTRCE 300
DB 241 FALTSLGDTLWTQWTRSIHACNKRGTGGRKEILSALYSPMDIQVLSQERQPFHTRCE 300
QY 301 EDNGGCSHLCLLSPEPYTCACPTGVQLQDNGRTCKAGAEVLLAARTDLRLISLDTTP 360
DB 301 EDNGGCSHLCLLSPEPYTCACPTGVQLQDNGRTCKAGAEVLLAARTDLRLISLDTTP 360

QY 361 DFTDIVLQVDDIRHAIADIDYDLEGVYVYVTTDDEVAIRRAIRRAYLDGSGAQLVNTINDEPDG 420
DB DFTDIVLQVDDIRHAIADIDYDLEGVYVYVTTDDEVAIRRAIRRAYLDGSGAQLVNTINDEPDG 420
QY 421 IAVDWARNLYWTDGTGTDRIEIVTRINGTSRKILVSIEDLDEPRALHHPVGLMYWTDNGE 480
DB IAVDWARNLYWTDGTGTDRIEIVTRINGTSRKILVSIEDLDEPRALHHPVGLMYWTDNGE 480
QY 481 NPKIECANLDGQERRVLVNASLGNPGLALDQSGKLYWGDAKTDKIIVNVVDGTKRRTL 540
DB NPKIECANLDGQERRVLVNASLGNPGLALDQSGKLYWGDAKTDKIIVNVVDGTKRRTL 540
QY 541 LEDKLPHIFGFTLLGDFIYWTDMQRRSRIERVHKVKSARDVIIDQLPDLMLGLKAVNAVAVV 600
DB LEDKLPHIFGFTLLGDFIYWTDMQRRSRIERVHKVKSARDVIIDQLPDLMLGLKAVNAVAVV 600
QY 601 GTNPCADRNGGCSHLCPFTPHATRCGPIGUELLSDMKTCIVPEAFVFTSRAAIIHRISL 660
DB GTNPCADRNGGCSHLCPFTPHATRCGPIGUELLSDMKTCIVPEAFVFTSRAAIIHRISL 660
QY 661 ETNNNDVAIPUTGVKEASALDPDVSNNHIIYWDVSLKTIISRAFMNGSSVEHVVEFGLDYP 720
DB ETNNNDVAIPUTGVKEASALDPDVSNNHIIYWDVSLKTIISRAFMNGSSVEHVVEFGLDYP 720
QY 721 EGMAVDMWGNKLYWADGTNRIEVARLDGQPRQVLVWRDLNPRSLALDPTKGYIYTWTEW 780
DB EGMAVDMWGNKLYWADGTNRIEVARLDGQPRQVLVWRDLNPRSLALDPTKGYIYTWTEW 780
QY 781 GKGPRIVRAFMGNTCMTLVDKVGANDLTIDYADQRLYWTDLDTNMISSNMLGQSRV 840
DB GKGPRIVRAFMGNTCMTLVDKVGANDLTIDYADQRLYWTDLDTNMISSNMLGQSRV 840
QY 841 IADLPHPGTGLTOYSDVIYWTDMNLHSTERADKTSGRNRTLIQGHLPFMDILVPHSSRQ 900
DB IADLPHPGTGLTOYSDVIYWTDMNLHSTERADKTSGRNRTLIQGHLPFMDILVPHSSRQ 900
QY 901 DGLNDCMHNNGCCQLCLAIIPGHRGCGASHYTLDPSSRNCSPPTTFLFQSKAISIRMI 960
DB DGLNDCMHNNGCCQLCLAIIPGHRGCGASHYTLDPSSRNCSPPTTFLFQSKAISIRMI 960
QY 961 PDDQHSPLILPLHGLRNVAIDYDPLDKFYIYWDGRQNIKRAKDDGTGPVFLSLSGQ 1020
DB PDDQHSPLILPLHGLRNVAIDYDPLDKFYIYWDGRQNIKRAKDDGTGPVFLSLSGQ 1020
QY 1021 NPDRQPHDLSIDIYSRTLFWTCEATNTINVHRLSSEAMGVVLRGDRDKPRAIVNAERG 1080
DB NPDRQPHDLSIDIYSRTLFWTCEATNTINVHRLSSEAMGVVLRGDRDKPRAIVNAERG 1080
QY 1081 LYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLFWVDADLKRIESCD 1140
DB LYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLFWVDADLKRIESCD 1140
QY 1141 LSGANRLTLEDANIVQPLGLTILGKHLIYIDRQOQMIERVEKTTGDKRTRIQGRVAHLTG 1200
DB LSGANRLTLEDANIVQPLGLTILGKHLIYIDRQOQMIERVEKTTGDKRTRIQGRVAHLTG 1200
QY 1201 IHAVEEVSLEFSAHPCARDNGGCSHICIAKDGCTPRCSCPVHLVLQNLITCCGEPPTCS 1260
DB IHAVEEVSLEFSAHPCARDNGGCSHICIAKDGCTPRCSCPVHLVLQNLITCCGEPPTCS 1260
QY 1261 PDQFACATGETIDCIIPGAWRCDDGPECCDDQSDDEECPCVCSAAQFPCARGQCVDLRLRCDE 1320
DB PDQFACATGETIDCIIPGAWRCDDGPECCDDQSDDEECPCVCSAAQFPCARGQCVDLRLRCDE 1320
QY 1321 ADCQDRSDEADCAICLPNPFRCASGCVLIIKQCDSPFDCIDGSDMLCMCITKPPSDDS 1380
DB ADCQDRSDEADCAICLPNPFRCASGCVLIIKQCDSPFDCIDGSDMLCMCITKPPSDDS 1380
QY 1381 PAHSSATGPVIGIITLSLFWMGVYVFCORVVCQRYAGANGPPHVEYVSGTTPHVLNFTAP 1440
DB PAHSSATGPVIGIITLSLFWMGVYVFCORVVCQRYAGANGPPHVEYVSGTTPHVLNFTAP 1440
QY 1441 GGSQHGPFTGIACGKSMWSSVSLMGGRGVPLYDRNHVTGASSSSSSSTKATLYPPIINP 1500

```

Db 1441 GSGQHGPTGIACGKSMSSVSLMGGGVLVDRNHTVGSSSSSTKATLYPILNP 1500
QY 1501 PPSPATDPSLVNMDYFSSNIPATARPYPYIIRGMWAPPTTPCSTDVCDSDYSASRWKAS 1560
Db 1501 PPSPATDPSLVNMDYFSSNIPATARPYPYIIRGMWAPPTTPCSTDVCDSDYSASRWKAS 1560
QY 1561 KYLDLNSDSDPYPPPTPHSOLSAEDSCPSPATERSYFHLFPPPSPTDSS 1615
Db 1561 KYLDLNSDSDPYPPPTPHSOLSAEDSCPSPATERSYFHLFPPPSPTDSS 1615

RESULT 3
ID ADB98798
XX ADB98798 standard; protein; 1615 AA.
AC ADB98798;
DT 04-DEC-2003 (first entry)
DE Human Zmax1(LRP5).
XX Osteopathic; Gene therapy; High Bone Mass; HBM; LRP5; Zmax1; LRP6;
KW bone mass modulation; osteoporosis; human.
OS Homo sapiens.
XX WO200292000-A2.
PD 21-NOV-2002.
XX 13-MAY-2002; 2002WO-US014877.
PR 11-MAY-2001; 2001US-0290071P.
PR 17-MAY-2001; 2001US-0291311P.
PR 01-FEB-2002; 2002US-0353058P.
PR 04-MAR-2002; 2002US-0361293P.
XX (GENO-) GENOME THERAPEUTICS CORP.
PA (AMHP ) WYETH.
XX
PI Allen K, Anisowicz A, Graham JR, Morales A, Yaworsky PJ, Liu W;
DR WPI; 2003-129214/12.
XX
PT New nucleic acid comprising a mutation in LRP5 or LRP6, useful for
PT diagnosing a HBM-like phenotype in a subject and for preparing a
PT composition for modulating bone mass and/or lipid levels in a subject
PT suffering from e.g. osteoporosis.
XX
PS Disclosure; Fig 26; 629pp; English.
XX
CC The present invention relates to High Bone Mass (HBM), LRP5 (Zmax1) and
CC LRP6 mutants, which results in a HBM-like phenotype when expressed in a
CC cell. The HBM-like phenotype results in bone mass modulation and/or lipid
CC level modulation. The invention is useful for diagnosing a HBM-like
CC phenotype in a subject and for preparing a composition for modulating
CC bone mass and/or lipid levels in a subject suffering from e.g.
CC osteoporosis. The present sequence was used to illustrate the invention.
XX
SQ Sequence 1615 AA;

Query Match 100.0%; Score 8740; DB 7; Length 1615;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1615; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEAAPGPPWPLLLLLLLLLLALCGCPAPAAAPLILFANRRDVRILVAGGVKLESTIVVS 60
Db 1 MEAAPGPPWPLLLLLLLLLLALCGCPAPAAAPLILFANRRDVRILVAGGVKLESTIVVS 60
QY 61 GLEDAAAVDFOFSKGVVWTVDSBEAIKQTYLNQTAGAQQVNVISGLVSPDGLACDWGK 120
Db 61 GLEDAAAVDFOFSKGVVWTVDSBEAIKQTYLNQTAGAQQVNVISGLVSPDGLACDWGK 120

```

```

QY 121 KLYWTDSETNRIEVANLNGTSRKVLFWQDLDPRAIALDPAHGWMYWTWGTETPRIERAG 180
Db 121 KLYWTDSETNRIEVANLNGTSRKVLFWQDLDPRAIALDPAHGWMYWTWGTETPRIERAG 180
QY 181 MDGSTRKIIIVDSDIYWPNGLTIDLEEOKLYWADAKLSPIHRANLDGSPROKVVESGLTHP 240
Db 181 MDGSTRKIIIVDSDIYWPNGLTIDLEEOKLYWADAKLSPIHRANLDGSPROKVVESGLTHP 240
QY 241 FALTSLSGDTLYWTDQWTRSIHACNKRITGGKKEILSALYSMDIQVLSQERQPPFHTRCE 300
Db 241 FALTSLSGDTLYWTDQWTRSIHACNKRITGGKKEILSALYSMDIQVLSQERQPPFHTRCE 300
QY 301 EDNGCCHLCLLSPSEPPYTCACPTGVLOQNGTCKAGAEVILLARTRDLRISLDTLP 360
Db 301 EDNGCCHLCLLSPSEPPYTCACPTGVLOQNGTCKAGAEVILLARTRDLRISLDTLP 360
QY 361 DFTDIVLQVDDIRHAIAIDYDPLEGYVYVWTDDEVRAIRRAYLDGSGAGTLVNTINDEPDG 420
Db 361 DFTDIVLQVDDIRHAIAIDYDPLEGYVYVWTDDEVRAIRRAYLDGSGAGTLVNTINDEPDG 420
QY 421 IAVDWARNLYWTDGTDRIEVTRNLNGTSRKILVSEDLDEPRAIALHPVGLMYWTDWGE 480
Db 421 IAVDWARNLYWTDGTDRIEVTRNLNGTSRKILVSEDLDEPRAIALHPVGLMYWTDWGE 480
QY 481 NPKTECANLDGQERRVLVNASLGNPGLALDLQEGKLYWGDATKIEVINVDGTKRRTL 540
Db 481 NPKTECANLDGQERRVLVNASLGNPGLALDLQEGKLYWGDATKIEVINVDGTKRRTL 540
QY 541 LEDKLPHIFGFTLLGDFIYWTDWQRRSIEVHVKASRDVIIDQLPDLMLKANVAVKV 600
Db 541 LEDKLPHIFGFTLLGDFIYWTDWQRRSIEVHVKASRDVIIDQLPDLMLKANVAVKV 600
QY 601 GTNPCADRNGGCSHLCPFTPHATRCGCPIGLELLSDMKTCIVPEAFVFTSRAAIHRISL 660
Db 601 GTNPCADRNGGCSHLCPFTPHATRCGCPIGLELLSDMKTCIVPEAFVFTSRAAIHRISL 660
QY 661 ETNNNDVAIPLTGVKEASALDFVSNHHIYWTDVSLAKTISRAFMNGSSVEHVFGDLYP 720
Db 661 ETNNNDVAIPLTGVKEASALDFVSNHHIYWTDVSLAKTISRAFMNGSSVEHVFGDLYP 720
QY 721 EGMADVMMGNKLYWADTGNRIEVARLDGQPROVLVWRDLNPRSLALDPTKGYIYWTEW 780
Db 721 EGMADVMMGNKLYWADTGNRIEVARLDGQPROVLVWRDLNPRSLALDPTKGYIYWTEW 780
QY 781 GKGPRIVRAFMWGNVCMTLVDKVRANDLTIDYADQRLYWTDLDTNMIESNNMLGQRRVV 840
Db 781 GKGPRIVRAFMWGNVCMTLVDKVRANDLTIDYADQRLYWTDLDTNMIESNNMLGQRRVV 840
QY 841 IADDLPHFPFGLTOYSDYIYWTDMNLHSTIRADKTSGRNRTLIOGHLDVMDILVFHSSRQ 900
Db 841 IADDLPHFPFGLTOYSDYIYWTDMNLHSTIRADKTSGRNRTLIOGHLDVMDILVFHSSRQ 900
QY 901 DGLMDCMHNNGQCQQLCIAIPGGHRCGCASHYTLDPSSRNCSPPTTFLFSQSAISRM 960
Db 901 DGLMDCMHNNGQCQQLCIAIPGGHRCGCASHYTLDPSSRNCSPPTTFLFSQSAISRM 960
QY 961 PDDQHSPLILPLHGLRNVAIDYDPLDKFYIYWDGQONIKRAKDDGTOPPVLTSLSQGQ 1020
Db 961 PDDQHSPLILPLHGLRNVAIDYDPLDKFYIYWDGQONIKRAKDDGTOPPVLTSLSQGQ 1020
QY 1021 NPDQPHDLSDIYSRRTLFWTCATNTINVHRLSGEAMGVVLRGDRDKPRAIVNAERGY 1080
Db 1021 NPDQPHDLSDIYSRRTLFWTCATNTINVHRLSGEAMGVVLRGDRDKPRAIVNAERGY 1080
QY 1081 LYFTNMDRAAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLFWVDADLAKRIESCD 1140
Db 1081 LYFTNMDRAAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLFWVDADLAKRIESCD 1140
QY 1141 LSGANRLTLEDANIVQPLGLTILGKHLVWDRQOQMIERVEKTTGDKRTRIOGRAVHLTG 1200
Db 1141 LSGANRLTLEDANIVQPLGLTILGKHLVWDRQOQMIERVEKTTGDKRTRIOGRAVHLTG 1200

```


Db 541 LEDKLPHIFGTLGDFIYWTWQRRSIEVHKVKSARDVIIDQLPDLMLKANVAKV 600
QY 601 GINPCADNRGGCSHLGFFTPHATRCGCPIGLELLSDMKTCIIVPEAFVFTSRAAIIHRLS 660
Db 601 GINPCADNRGGCSHLGFFTPHATRCGCPIGLELLSDMKTCIIVPEAFVFTSRAAIIHRLS 660
QY 661 ETNNNDVAIPLTGVKEASALDFDVSNHIIYWTDSIKTISRPFMGSSVEHVVEGLDYP 720
Db 661 ETNNNDVAIPLTGVKEASALDFDVSNHIIYWTDSIKTISRPFMGSSVEHVVEGLDYP 720
QY 721 EGMAYDWMGKNLYADTGTNRLEVARLDGQFQVLRDLNPRSLALDPTKGYIYWTW 780
Db 721 EGMAYDWMGKNLYADTGTNRLEVARLDGQFQVLRDLNPRSLALDPTKGYIYWTW 780
QY 781 GSKPRIVAFMDGTNCMTLVKVRANDLTIDYADQRLYWTDLDTNMIESSNMLGQERV 840
Db 781 GSKPRIVAFMDGTNCMTLVKVRANDLTIDYADQRLYWTDLDTNMIESSNMLGQERV 840
QY 841 IADDLPHPEGLTOYSDYIYWTWNHLSITERADKTSGRNRTLIQHLDFVMDILVPHSSRQ 900
Db 841 IADDLPHPEGLTOYSDYIYWTWNHLSITERADKTSGRNRTLIQHLDFVMDILVPHSSRQ 900
QY 901 DGLNDGMHNGCGQLCLAI PGHRCGCASHYTLDPSSRNCSPPTTFLFSOKSAISMI 960
Db 901 DGLNDGMHNGCGQLCLAI PGHRCGCASHYTLDPSSRNCSPPTTFLFSOKSAISMI 960
QY 961 PDDQHSPLILPLHGLRNKAIIDYDPLDKFIYWDGRQNIKRAKDDGTQPPVLTSLSGQ 1020
Db 961 PDDQHSPLILPLHGLRNKAIIDYDPLDKFIYWDGRQNIKRAKDDGTQPPVLTSLSGQ 1020
QY 1021 NPDROPHDLSIDIYBRTLFTWCEANTINVELSGEAMGVLRGDRDKPRAIVNAERG 1080
Db 1021 NPDROPHDLSIDIYBRTLFTWCEANTINVELSGEAMGVLRGDRDKPRAIVNAERG 1080
QY 1081 LYFTWQDRAAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLKFWVDADLKRIESC 1140
Db 1081 LYFTWQDRAAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLKFWVDADLKRIESC 1140
QY 1141 LSGANRLTLEDANIYVQILGLTILGHLYWIDRQQMIERVEKTDGKTRIQGRVAHLTG 1200
Db 1141 LSGANRLTLEDANIYVQILGLTILGHLYWIDRQQMIERVEKTDGKTRIQGRVAHLTG 1200
QY 1201 THAVEEVSLEFSAHPCARDNGGCSHICIAKGDTGPRCSPVHLVLLQNLATCGEPPTCS 1260
Db 1201 THAVEEVSLEFSAHPCARDNGGCSHICIAKGDTGPRCSPVHLVLLQNLATCGEPPTCS 1260
QY 1261 PDQFACATGEIDCIPGAWRCDFPECDQSBEGCFVCSAAQFPFCARQCQVDLRLCDGE 1320
Db 1261 PDQFACATGEIDCIPGAWRCDFPECDQSBEGCFVCSAAQFPFCARQCQVDLRLCDGE 1320
QY 1321 ADCQDRSEADCDALCLPNQRCASGQVLIKQCCDFFDCIDGSDLMCEITPPSDDS 1380
Db 1321 ADCQDRSEADCDALCLPNQRCASGQVLIKQCCDFFDCIDGSDLMCEITPPSDDS 1380
QY 1381 PAHSSAIGFVIGIILSLFVMGCVFVFCORVQVAGANGPPFHYVSGTVPVNFATP 1440
Db 1381 PAHSSAIGFVIGIILSLFVMGCVFVFCORVQVAGANGPPFHYVSGTVPVNFATP 1440
QY 1441 GGSQHGPTGTIACGKSMSSVSLMGGRGGVPLYDRNHVTGASSSSSSSTKATLYPPLNP 1500
Db 1441 GGSQHGPTGTIACGKSMSSVSLMGGRGGVPLYDRNHVTGASSSSSSSTKATLYPPLNP 1500
QY 1501 PPSPATDSLYNMDFYSSNIPATAPRYPYIIRGMAPPPTPCSDVCDSDYSASRWKAS 1560
Db 1501 PPSPATDSLYNMDFYSSNIPATAPRYPYIIRGMAPPPTPCSDVCDSDYSASRWKAS 1560
QY 1561 KYIYLDNDSDPYPPPTPHSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615
Db 1561 KYIYLDNDSDPYPPPTPHSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615

AAW83309
ID AAW83309 standard; protein; 1615 AA.
XX
AC AAW83309;
XX
DT 10-FEB-1999 (first entry)
XX
DE LRP5 protein from the longest open reading frame.
XX
KW LRP5; LDL-receptor related protein; LRP-3; IDDM; diagnosis; endocytosis;
KW insulin dependent diabetes mellitus; autoimmune disease;
KW glomerulonephritis; inflammation; viral infection; osteoporosis;
KW hypercholesterolemia; Alzheimer's disease; low density lipoprotein.
XX
OS Homo sapiens.
XX
PN WO9846743-A1.
XX
PD 22-OCT-1998.
XX
PF 15-APR-1998; 98WO-GB001102.
XX
PR 15-APR-1997; 97US-0043553P.
PR 05-JUN-1997; 97US-0048740P.
XX
PA (WELL) WELLCOME TRUST LTD.
PA (MERI) MERCK & CO INC.
XX
XX Todd JA, Hess JW, Caskey CT, Cox RD, Gerhold D, Hammond H;
PI Hey P, Kawaguchi Y, Merriman TR, Metzker ML, Nakagawa Y;
PI Phillips MS, Twells RCU;
XX
DR WPI; 1998-594573/50.
XX
XX New isolated LDL-receptor related protein - used to develop products for
PT treating, e.g. elevated triglyceride levels, diabetes, autoimmune
PT disorders, inflammation or Alzheimer's disease.
XX
PS Claim 2; Fig 5c; 200pp; English.

The present sequence represents a LRP5 protein (low density lipoprotein
(LDL) receptor related protein, previously designated LRP-3). Nucleic
acid molecules (NAMS) encoding LRP5 can be used for determining if an
individual is susceptible to insulin dependent diabetes mellitus (IDDM).
The NAMS or proteins can be used for reducing triglyceride levels in the
serum of an individual. Therapies that affect LRP5 may also be useful in
the treatment of autoimmune diseases such as glomerulonephritis, diseases
and disorders involving disruption of endocytosis and/or antigen
presentation, cytokine clearance and/or inflammation, viral infection,
pathogenic bacterial toxin contamination, elevation of free fatty acids
or hypercholesterolemia, type 2 diabetes, osteoporosis, Alzheimer's
disease and cardiovascular disease. Products from the present invention
can also be used for detection, diagnosis and drug screening

Sequence 1615 AA;

Query Match 100.0%; Score 8736; DB 2; Length 1615;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1614; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MEAAPPPPPPLLLLLLLLLLALCGCPAPAAASPLLLFANRRDVRVLVDAGGVKLESTIVS 60
Db 1 MEAAPPPPPPLLLLLLLLLLALCGCPAPAAASPLLLFANRRDVRVLVDAGGVKLESTIVS 60
QY 61 GLEDAAAVDVDFQSKGAVYWTVDVSEAIKQTYLNTQTAQVNVVTSGLVSPDGLADWVGK 120
Db 61 GLEDAAAVDVDFQSKGAVYWTVDVSEAIKQTYLNTQTAQVNVVTSGLVSPDGLADWVGK 120
QY 121 KLYWTDSETNRIEVANLNGTSRKVLFWQDDQRAIALDPAHGYWYTDGSETPRIERAG 180
Db 121 KLYWTDSETNRIEVANLNGTSRKVLFWQDDQRAIALDPAHGYWYTDGSETPRIERAG 180
QY 181 MDGSTRKIIIVDSDIYWPNGLTIDLEQKLYWADAKLSFIHRANLDGSRQKRVVSGSLTHP 240

Db MDGSTRKIIIVDSIYWPNGLTIDLEQKLYWADAKLSPIHRANLDGSRQKRVESGLTLP 240
QY FALTLSGDTLTYWTDWQTSIIHACNKRITGGKKEILSALYSPMDIQVLQSRQPPFHTRCE 300
Db FALTLSGDTLTYWTDWQTSIIHACNKRITGGKKEILSALYSPMDIQVLQSRQPPFHTRCE 300
QY EDNGGCSHLCLLSPEPEYTCACPTGVQIQNGRTCKAGAEVLLIARRTLRLRLSLDTP 360
Db EDNGGCSHLCLLSPEPEYTCACPTGVQIQNGRTCKAGAEVLLIARRTLRLRLSLDTP 360
QY DFTDVLQVDDIRHAIADYDPLEGYVYWTDEVRRAIRRAYLDGSGAQTLVNTEINDPDG 420
Db DFTDVLQVDDIRHAIADYDPLEGYVYWTDEVRRAIRRAYLDGSGAQTLVNTEINDPDG 420
QY IAVDWARNLYWTDGTDRIEVTRLANGTSRKILVSEDLDEPRATLHFPVGLMTWDGGE 480
Db IAVDWARNLYWTDGTDRIEVTRLANGTSRKILVSEDLDEPRATLHFPVGLMTWDGGE 480
QY NPKIECANLDQERRVLYNVAISLGNPGLALDLQEGKLYWDAKTDKIEVINVDGTKRRTL 540
Db NPKIECANLDQERRVLYNVAISLGNPGLALDLQEGKLYWDAKTDKIEVINVDGTKRRTL 540
QY LEDKLPHPFGTLLGDFIYWTDMQRRSIEVHKVKAASRDVIIDQLPDLMLGLKAVNAVAV 600
Db LEDKLPHPFGTLLGDFIYWTDMQRRSIEVHKVKAASRDVIIDQLPDLMLGLKAVNAVAV 600
QY GTNPNACDRNGCSHLCPFTPHATRCGCPIGIELISDMKTCIVPEAFVFTSRAAIIHRLS 660
Db GTNPNACDRNGCSHLCPFTPHATRCGCPIGIELISDMKTCIVPEAFVFTSRAAIIHRLS 660
QY ETNNNDVAIPLTGVKEASALDPDVSNHHIYWTDSVLSKTIISAFMNGSSVEHVFEGLDYP 720
Db ETNNNDVAIPLTGVKEASALDPDVSNHHIYWTDSVLSKTIISAFMNGSSVEHVFEGLDYP 720
QY EGMVAVDMGKGLYNWADTGTNRIEVARLDGQFQVLRVDRDLNPRSLALDPTKGIYIYTW 780
Db EGMVAVDMGKGLYNWADTGTNRIEVARLDGQFQVLRVDRDLNPRSLALDPTKGIYIYTW 780
QY GKGPRIVRAFDGNTCNMTLVKVGANDLTIDYADQRLYWTDLDTNMISSNMLQGERV 840
Db GKGPRIVRAFDGNTCNMTLVKVGANDLTIDYADQRLYWTDLDTNMISSNMLQGERV 840
QY IADLLPHFPGLTQYSDIYWTDMNLHSIERADKTSGRNRTLIQGHLDFFWMDILVPHSSRQ 900
Db IADLLPHFPGLTQYSDIYWTDMNLHSIERADKTSGRNRTLIQGHLDFFWMDILVPHSSRQ 900
QY DGLNDCMHNNGCQCGLCLAI PGHRCGCASHYTLDPSSRNCSPPTTFLPSQSAISRMI 960
Db DGLNDCMHNNGCQCGLCLAI PGHRCGCASHYTLDPSSRNCSPPTTFLPSQSAISRMI 960
QY PDDQHSPLIILPLHGLRNVAIDYDPLDKFYIYWDGQNIKRAKDDGTQPPVLISLQGG 1020
Db PDDQHSPLIILPLHGLRNVAIDYDPLDKFYIYWDGQNIKRAKDDGTQPPVLISLQGG 1020
QY NPDRQPHDLSIDIYSRITLFWTCEATNTINVRHLSGEAMGVVLRGDRDKPRAIVNAERGY 1080
Db NPDRQPHDLSIDIYSRITLFWTCEATNTINVRHLSGEAMGVVLRGDRDKPRAIVNAERGY 1080
QY LYFTNMQDRAAKIERAALDGTGEREVLTFTGLIRPVVALVVDNTLGLKFWVDADLKRIESED 1140
Db LYFTNMQDRAAKIERAALDGTGEREVLTFTGLIRPVVALVVDNTLGLKFWVDADLKRIESED 1140
QY LSGANRLTLEANIYVQPLGLTILGKHLWIDRQOQMIERVEKTTGDKRTRIOGRAVHLTG 1200
Db LSGANRLTLEANIYVQPLGLTILGKHLWIDRQOQMIERVEKTTGDKRTRIOGRAVHLTG 1200
QY IHAVEEVSLEFSAHPCARDNGGCSHICIAKGDGTPRCSVPHLVLQNLTLTCEBPPTCS 1260
Db IHAVEEVSLEFSAHPCARDNGGCSHICIAKGDGTPRCSVPHLVLQNLTLTCEBPPTCS 1260
QY PDQFACATGEIDCIPGAWRCDFPECDQSDGCPVCSAAQFFPCARQCQVDRLRLRCDGE 1320

Db PDQFACATGEIDCIPGAWRCDFPECDQSDGCPVCSAAQFFPCARQCQVDRLRLRCDGE 1320
QY ADCQDRSDEADCAICLPNQFRCSAGQCVLIKQCCDSFPDCIDGSDLMCMCITKPPSDDS 1380
Db ADCQDRSDEADCAICLPNQFRCSAGQCVLIKQCCDSFPDCIDGSDLMCMCITKPPSDDS 1380
QY PAHSSAIGPVIIGIILSLFVMGGVYFVQVVCQRYAGANGFPPEHYVSGTTPHVLNFIAP 1440
Db PAHSSAIGPVIIGIILSLFVMGGVYFVQVVCQRYAGANGFPPEHYVSGTTPHVLNFIAP 1440
QY GGSQHGFPTGIACGKSMSSVSLMGGRGVPLYDRNHVHTGASSSSSSSTKATLYPILNP 1500
Db GGSQHGFPTGIACGKSMSSVSLMGGRGVPLYDRNHVHTGASSSSSSSTKATLYPILNP 1500
QY PPSPATPSLYNMDMFYSSNIPATRPVRYPIIRGMAPPPTPCSTDVDCSDYSASRWKAS 1560
Db PPSPATPSLYNMDMFYSSNIPATRPVRYPIIRGMAPPPTPCSTDVDCSDYSASRWKAS 1560
QY KYLLDLNSDSDPYPPPTPHSQYLSAEDSCPPSPATERSYFHLFPPPPSPCTDSS 1615
Db KYLLDLNSDSDPYPPPTPHSQYLSAEDSCPPSPATERSYFHLFPPPPSPCTDSS 1615

RESULT 6

AAE21740

ID AAE21740 standard; protein; 1615 AA.

XX AC AAE21740;

XX DT 16-JUL-2002 (first entry)

XX DE Human BSMR protein mutant, R494Q.

XX KW Human; bone strength and mineralisation regulatory protein; BSMR;
bone strength; mineralisation; ophthalmological; antidiabetic;
bone density regulating transmembrane receptor; prosthetic device;
surgical implant; diabetic retinopathy; hypertensive retinopathy;
therapy; osteoporosis; prematurity; ocular vessel; eye disorder;
osteopathic; mutant; mutein.

XX OS Homo sapiens.

XX OS Synthetic.

XX FH Key Location/Qualifiers

XX FT Misc-difference 494

XX FT /note= "Wild-type Arg substituted with Gln"

XX WO200216553-A2.

XX PD 28-FEB-2002.

XX PF 17-AUG-2001; 2001WO-US041788.

XX PR 18-AUG-2000; 2000US-0226119P.

XX PR 22-SEP-2000; 2000US-0234337P.

XX PR 13-JUL-2001; 2001US-0304851P.

XX PA (AVET) AVENTIS PHARMA SA.

XX PA (HARD) HARVARD COLLEGE.

XX PA (UYCA-) UNIV CASE WESTERN RESERVE.

XX PI Warman ML, Gong Y, Olsen BR, Rawadi G, Roman-Roman S;

XX DR WPI; 2002-329694/36.

XX PT Polynucleotide encoding bone strength and mineralization regulatory

XX PT protein useful for diagnosis or therapy of osteoporosis.

XX PS Disclosure; Page; 124pp; English.

XX CC The invention relates to bone strength and mineralisation regulatory

XX CC protein (BSMR) and its corresponding nucleic acid sequence. BSMR DNA is

XX CC useful for the diagnosis or therapy of osteoporosis and for regulating

(increasing) bone strength and mineralisation in a human subject by activating a bone density regulating transmembrane receptor (BSMR protein). An expression vector comprising a promoter that is operably linked to BSMR DNA is useful for modulating bone density and for enhancing bone strength and mineralisation in a mammal cell. Composition comprising a BSMR effector is useful for treating osteoporosis and is useful particularly as a coating for prosthetic devices and surgical implants. BSMR is useful for screening lead pharmaceutical agents as BSMR effectors, which may be used to treat a range of eye disorders such as diabetic retinopathy, hypertensive retinopathy and retinopathy of prematurity, in which normal vascular growth and integrity of ocular vessels is disrupted. The present sequence is human BSMR protein mutant. Note: This sequence is not shown in the specification, however it is constructed based on the human BSMR protein shown in fig 3 of the specification (AAE21730)

XX SQ Sequence 1615 AA;

Query Match 100.0%; Score 8736; DB 5; Length 1615;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1614; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEAAPGGPPWPLLLLLLLLLLALCGCPAPAAAPLLLFANRRDRLVLDAGGVKLESTIVVS 60
 DB 1 MEAAPGGPPWPLLLLLLLLLLALCGCPAPAAAPLLLFANRRDRLVLDAGGVKLESTIVVS 60

QY 61 GLEDAADVDFQSKGAVYWDVSEAIKQTYLNQTAAGVONVVISGLVSPDGLACDWGK 120
 DB 61 GLEDAADVDFQSKGAVYWDVSEAIKQTYLNQTAAGVONVVISGLVSPDGLACDWGK 120

QY 121 KLYWTDSETNRELVANLNGTSRKVLFWQDLDOPRAIALDPAGHYMYWTDWGETPRIERAG 180
 DB 121 KLYWTDSETNRELVANLNGTSRKVLFWQDLDOPRAIALDPAGHYMYWTDWGETPRIERAG 180

QY 181 MDGSTRKLIIVSDIYWPNGLTIDLEQKLYWADAKLSFIHRANLDGSPRQKVVEGSLTHP 240
 DB 181 MDGSTRKLIIVSDIYWPNGLTIDLEQKLYWADAKLSFIHRANLDGSPRQKVVEGSLTHP 240

QY 241 FALTISGDTLYWTDQWTSIHACNKRGTGKRKEILSALYSPMDIQLVQERQPPFHTRCE 300
 DB 241 FALTISGDTLYWTDQWTSIHACNKRGTGKRKEILSALYSPMDIQLVQERQPPFHTRCE 300

QY 301 EDNGGCSHLCLLSPSEPFYTCPTGVQLQDNGRTCKAGAEVILLARRTLRLRLSLDTP 360
 DB 301 EDNGGCSHLCLLSPSEPFYTCPTGVQLQDNGRTCKAGAEVILLARRTLRLRLSLDTP 360

QY 361 DPTDVLQVDDIRHAIAIDYDPLEGYVYWDDEVRAIRRAYLDGSGAQLVNTENDPDG 420
 DB 361 DPTDVLQVDDIRHAIAIDYDPLEGYVYWDDEVRAIRRAYLDGSGAQLVNTENDPDG 420

QY 421 IAVDVARNLYWTDGTDRIEVTRNLNGTSRKILVSEDLDEPRALHFPVGLMYWTDWGE 480
 DB 421 IAVDVARNLYWTDGTDRIEVTRNLNGTSRKILVSEDLDEPRALHFPVGLMYWTDWGE 480

QY 481 NPKIECANLDGQERRVLVNASLWPNGLALDQEGKLYWGAQTKDIEVINVDGTKRRTL 540
 DB 481 NPKIECANLDGQERRVLVNASLWPNGLALDQEGKLYWGAQTKDIEVINVDGTKRRTL 540

QY 541 LEDKLPHIFGFTLLGDFIYWDQRRSIERVHKVKAASRDVITDQLPDLMLGLKAVNAVAVV 600
 DB 541 LEDKLPHIFGFTLLGDFIYWDQRRSIERVHKVKAASRDVITDQLPDLMLGLKAVNAVAVV 600

QY 601 GTNPCADNRGGCSHLCTFTPHATRCGCPIGLELLSDMKTICVPEAFVFTSRAIHRISL 660
 DB 601 GTNPCADNRGGCSHLCTFTPHATRCGCPIGLELLSDMKTICVPEAFVFTSRAIHRISL 660

QY 661 ETNNNDVAIPTGVKEASALDFDVSNHHIYWDVSLKTIISRAFMNGSSVEHWFEGLDYP 720
 DB 661 ETNNNDVAIPTGVKEASALDFDVSNHHIYWDVSLKTIISRAFMNGSSVEHWFEGLDYP 720

QY 721 EGMAYDWMGKLYWADTGTNRIEVARLDGQFQVLVWRDLNPRSLALDPTKGYIYWTEW 780
 DB 721 EGMAYDWMGKLYWADTGTNRIEVARLDGQFQVLVWRDLNPRSLALDPTKGYIYWTEW 780

QY 781 GKGPRIVRAFAFGNVCMTLVKVGANDLTIDYADQRLYTDLDTNMTIENSNMLQERVV 840
 DB 781 GKGPRIVRAFAFGNVCMTLVKVGANDLTIDYADQRLYTDLDTNMTIENSNMLQERVV 840

QY 841 IADDLPHFPGTQVSDYIYWDNHLSTERADKTSGRNRTLIQGHLDPMWDLVPHSSRQ 900
 DB 841 IADDLPHFPGTQVSDYIYWDNHLSTERADKTSGRNRTLIQGHLDPMWDLVPHSSRQ 900

QY 901 DGLNDCMHNQCQGLCLAIPEGHRCGCASHYTLDPSSRNCSPPTTFLLFQSKAISMI 960
 DB 901 DGLNDCMHNQCQGLCLAIPEGHRCGCASHYTLDPSSRNCSPPTTFLLFQSKAISMI 960

QY 961 PDDQHSPLILPLHGLRNVAIDYDPLDKFIYWDGRQNIKRAKDXDGTQPFVLTSLSQGG 1020
 DB 961 PDDQHSPLILPLHGLRNVAIDYDPLDKFIYWDGRQNIKRAKDXDGTQPFVLTSLSQGG 1020

QY 1021 NPDRQPHDLSIDIYSRTLFWTCEATNTINVRHLSGEAMGVVLGRDOKPRAIVVNAERG 1080
 DB 1021 NPDRQPHDLSIDIYSRTLFWTCEATNTINVRHLSGEAMGVVLGRDOKPRAIVVNAERG 1080

QY 1081 LYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLFWVDADLKRIESCD 1140
 DB 1081 LYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLFWVDADLKRIESCD 1140

QY 1141 LSGANRLTLEDANIVQPLGLTILGKHLIYWDROQOMIERVEKTTGDKRTRIQGRVAHLTG 1200
 DB 1141 LSGANRLTLEDANIVQPLGLTILGKHLIYWDROQOMIERVEKTTGDKRTRIQGRVAHLTG 1200

QY 1201 IHAVEEVSLEEFSAHPCARDNGGCSHICIAKDGCTPRCSCPVHLVLLQNLATCCGEPPTCS 1260
 DB 1201 IHAVEEVSLEEFSAHPCARDNGGCSHICIAKDGCTPRCSCPVHLVLLQNLATCCGEPPTCS 1260

QY 1261 PDQFACATGEIDCIPGAWRCDFPECDQSDDEGCPVCSAAQFFCARGQCVDLRLCDGE 1320
 DB 1261 PDQFACATGEIDCIPGAWRCDFPECDQSDDEGCPVCSAAQFFCARGQCVDLRLCDGE 1320

QY 1321 ADCQDRSEADCAICLPNQPRCAGQCCLIKQCCDSFPDCIDGSDMLMCEBITKPPSDDS 1380
 DB 1321 ADCQDRSEADCAICLPNQPRCAGQCCLIKQCCDSFPDCIDGSDMLMCEBITKPPSDDS 1380

QY 1381 PAHSSAIGPVGIIISLFLVMGGVYFVQVVCQRYAGANGPFHEYVSGTTPHVLNFIAP 1440
 DB 1381 PAHSSAIGPVGIIISLFLVMGGVYFVQVVCQRYAGANGPFHEYVSGTTPHVLNFIAP 1440

QY 1441 GGSQHGPFPTGIACCKSMWSSVSLMGGRGVPLYDRNHVTGASSSSSSSTKATLYPPIINP 1500
 DB 1441 GGSQHGPFPTGIACCKSMWSSVSLMGGRGVPLYDRNHVTGASSSSSSSTKATLYPPIINP 1500

QY 1501 PPSPATPSLYNMDFYSSNIPATARYPIYIRGMAPPPTPCSTDYCDSDYSASRWKAS 1560
 DB 1501 PPSPATPSLYNMDFYSSNIPATARYPIYIRGMAPPPTPCSTDYCDSDYSASRWKAS 1560

QY 1561 KYILDLNSDSDPYPPPTPHSQVLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615
 DB 1561 KYILDLNSDSDPYPPPTPHSQVLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615

RESULT 7
 ADI27181
 ID ADI27181 standard; protein; 1615 AA.
 AC ADI27181;
 XX
 DT 22-APR-2004 (first entry)
 XX
 DE Human LRP binding family protein #12.
 XX
 KW tooth development; Wnt pathway activation; bone deposition;
 XX ocular development; cysteine knot; LRP binding family; human.
 OS Homo sapiens.
 XX

PN WO2003106657-A2.
 XX 24-DEC-2003.
 PD 16-JUN-2003; 2003WO-US019260.
 PF 14-JUN-2002; 2002US-0388970P.
 PR (STOW-) STOWERS INST MEDICAL RES.
 XX Krumlauf R, Ellies D;
 XX WPI; 2004-082189/08.
 XX Family of polypeptides having amino acid sequences that influence tooth
 PT development, Wnt pathway activation, bone deposition, and/or ocular
 PT development.
 XX Claim 126; SEQ ID NO 82; 517pp; English.
 XX The invention relates to a family of nucleic acid molecules which can
 CC influence tooth development, Wnt pathway activation, bone deposition or
 CC ocular development: expressing cysteine knot protein. The nucleic acids
 CC and the proteins are useful for increasing bone deposition, affecting the
 CC Wnt pathway, affecting tooth development, affecting ocular development
 CC and predicting defect in bone deposition. The present sequence represents
 CC the amino acid sequence of a LRP binding family protein.
 XX Sequence 1615 AA;
 SQ
 Query Match 100.0%; Score 8736; DB 8; Length 1615;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1614; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MEAAPGGPFWLLLLLLLLALCGCPAPAAASPLLLPANRRDRLVLDAGGVKLESTIVVS 60
 DB 1 MEAAPGGPFWLLLLLLLLALCGCPAPAAASPLLLPANRRDRLVLDAGGVKLESTIVVS 60
 QY 61 GLEDAADVDFPFGKAVYTWDSSEAIKQTYLNOTGAAVQNVVLSGLVSPDLACDWGK 120
 DB 61 GLEDAADVDFPFGKAVYTWDSSEAIKQTYLNOTGAAVQNVVLSGLVSPDLACDWGK 120
 QY 121 KLYWTDSTNRLEVANLNGTSRKVLFWQDLQPRALALDPAGHYMYTWDGETPRIERAG 180
 DB 121 KLYWTDSTNRLEVANLNGTSRKVLFWQDLQPRALALDPAGHYMYTWDGETPRIERAG 180
 QY 181 MDGSTRKLIIVDSIYWPNGLIIDLEQKLYWADAKLSFIHRANLDGSPRQKRVGSLTHP 240
 DB 181 MDGSTRKLIIVDSIYWPNGLIIDLEQKLYWADAKLSFIHRANLDGSPRQKRVGSLTHP 240
 QY 241 FALTLSGDTLYWTDQTRS IHACNKRKTGKKEILSALYSMDIQVLSQERQPPFHTTCE 300
 DB 241 FALTLSGDTLYWTDQTRS IHACNKRKTGKKEILSALYSMDIQVLSQERQPPFHTTCE 300
 QY 301 EDNGSCSHLCLLSPSEPFYTCACPTGVQLQNGRTCKAGABEVLLARTRDLRISLDP 360
 DB 301 EDNGSCSHLCLLSPSEPFYTCACPTGVQLQNGRTCKAGABEVLLARTRDLRISLDP 360
 QY 361 DFTDVLQVDDIRHAIAIDYDPLEGYVYTWDDVRAIRRAYLDGSGAOTLVNTEINDPDG 420
 DB 361 DFTDVLQVDDIRHAIAIDYDPLEGYVYTWDDVRAIRRAYLDGSGAOTLVNTEINDPDG 420
 QY 421 IAVDWARNLYWTDGTDR IEVTRNGTSRKILYSEDLDEPRALHPVGMGLMYTWDMGE 480
 DB 421 IAVDWARNLYWTDGTDR IEVTRNGTSRKILYSEDLDEPRALHPVGMGLMYTWDMGE 480
 QY 481 NPKIECANLQGERVLVNASIGWENGLALDLQEGKLYWDAKTDKIEVINVDGTKRTL 540
 DB 481 NPKIECANLQGERVLVNASIGWENGLALDLQEGKLYWDAKTDKIEVINVDGTKRTL 540
 QY 541 LEDKLPHPFGFTLLGDFYTWDMQRRS IERHVHVKASRDVIIDQLPDLMLGLKAVNAKV 600
 DB 541 LEDKLPHPFGFTLLGDFYTWDMQRRS IERHVHVKASRDVIIDQLPDLMLGLKAVNAKV 600

RESULT 8
 ADQ20524
 ID ADQ20524 standard; protein; 1615 AA.

QY 601 GTNPCADNRNGCCHLCFFTPHATRCGCPGLELLSDMKTCIVPEAFVFTSRAAIIHRISL 660
 DB 601 GTNPCADNRNGCCHLCFFTPHATRCGCPGLELLSDMKTCIVPEAFVFTSRAAIIHRISL 660
 QY 661 ETNNNDVAIPLTGVKEASALDFDVSNHHIYMTDVSLSKTIISRAFMNGSSVEHVVFGLDYP 720
 DB 661 ETNNNDVAIPLTGVKEASALDFDVSNHHIYMTDVSLSKTIISRAFMNGSSVEHVVFGLDYP 720
 QY 721 EGMADVMMGNKLYWADTGTNRIEVARLDGQROVLVWRDLNPRSLADPTKGYIYWTEW 780
 DB 721 EGMADVMMGNKLYWADTGTNRIEVARLDGQROVLVWRDLNPRSLADPTKGYIYWTEW 780
 QY 781 GGGKRIIVRAFAFDGNCMTLVKVGGRANDLTIDYADQRLYWTDLDTNMTLESSNMLQSERV 840
 DB 781 GGGKRIIVRAFAFDGNCMTLVKVGGRANDLTIDYADQRLYWTDLDTNMTLESSNMLQSERV 840
 QY 841 IADDLPHFPGLTQYSDIYIYWTDMNLHLSIERADKTSGRNRTLIQGHLDVFMDILVFHSSRQ 900
 DB 841 IADDLPHFPGLTQYSDIYIYWTDMNLHLSIERADKTSGRNRTLIQGHLDVFMDILVFHSSRQ 900
 QY 901 DGLNDCHNNGCQCGOLCLAIPEGHRCGCASHYTLDPSRNCSPPTTFLFSQKSAISRMI 960
 DB 901 DGLNDCHNNGCQCGOLCLAIPEGHRCGCASHYTLDPSRNCSPPTTFLFSQKSAISRMI 960
 QY 961 PDDQHSPLIILPLHGLRNKVIDYDPLDKFYIYWDGRQNIKRAKDDGTQPFVLTSLSQG 1020
 DB 961 PDDQHSPLIILPLHGLRNKVIDYDPLDKFYIYWDGRQNIKRAKDDGTQPFVLTSLSQG 1020
 QY 1021 NPDRQPHDLSIDIYSRTLFWTCEATNTINVHRLGSEAMGVVLRGDRDKPRAIVVNAERG 1080
 DB 1021 NPDRQPHDLSIDIYSRTLFWTCEATNTINVHRLGSEAMGVVLRGDRDKPRAIVVNAERG 1080
 QY 1081 LYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLKFWVDADLKIESCD 1140
 DB 1081 LYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLKFWVDADLKIESCD 1140
 QY 1141 LSGANRLTLEDANIVQPLGLTILGKHLWIQOQMIERVEKTTGDKETRIQGRVAHLTG 1200
 DB 1141 LSGANRLTLEDANIVQPLGLTILGKHLWIQOQMIERVEKTTGDKETRIQGRVAHLTG 1200
 QY 1201 IHAVEVSLEEFSAHPCARDNGGCSHICIAKDGTPRCSCPVLVLLQNLTLTCSBPPTCS 1260
 DB 1201 IHAVEVSLEEFSAHPCARDNGGCSHICIAKDGTPRCSCPVLVLLQNLTLTCSBPPTCS 1260
 QY 1261 PDQFACATGEIDCIPGAWRCDFPECDQSDDEGCPVCSAAQFPCCARGCQVDLRUCDGE 1320
 DB 1261 PDQFACATGEIDCIPGAWRCDFPECDQSDDEGCPVCSAAQFPCCARGCQVDLRUCDGE 1320
 QY 1321 ADCQDRSDEADCDALCLPNQPCASGOCVLLIKQCDSPDCIDGSDLMCEITKPPSDDS 1380
 DB 1321 ADCQDRSDEADCDALCLPNQPCASGOCVLLIKQCDSPDCIDGSDLMCEITKPPSDDS 1380
 QY 1381 PAHSSAIGPVGIIILSLFVMGVYFVQVRVQVRVQVRVQVRVQVRVQVRVQVRVQVRV 1440
 DB 1381 PAHSSAIGPVGIIILSLFVMGVYFVQVRVQVRVQVRVQVRVQVRVQVRVQVRV 1440
 QY 1441 GGSQHGPTGTIACGKSMSSVSLMGGRGVPLVDRNHVGTGASSSSSSSTKATLPPILNP 1500
 DB 1441 GGSQHGPTGTIACGKSMSSVSLMGGRGVPLVDRNHVGTGASSSSSSSTKATLPPILNP 1500
 QY 1501 PPSPATPSLVNMDMFYSSNIPATVPYRIIRGMAPPTTCTDVCDSYASRWKAS 1560
 DB 1501 PPSPATPSLVNMDMFYSSNIPATVPYRIIRGMAPPTTCTDVCDSYASRWKAS 1560
 QY 1561 KYIYDLNSDSDPYPPPTPHPSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615
 DB 1561 KYIYDLNSDSDPYPPPTPHPSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615

XX AC ADQ20524;
XX DT 26-AUG-2004 (first entry)
XX DE Human soft tissue sarcoma-upregulated protein - SEQ ID 3344.
XX KW soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human.
XX OS Homo sapiens.
XX PN WO2004048938-A2.
XX PD 10-JUN-2004.
XX PF 26-NOV-2003; 2003WO-US038193.
XX PR 26-NOV-2002; 2002US-0429739P.
XX PA (PROT-) PROTEIN DESIGN LABS INC.
XX PI Aziz N, Ginsburg WM, Zlotnik A;
XX DR WPI; 2004-441208/41.
XX PT Early detection of soft tissue sarcoma comprises determining expression
XX PT of a gene in a first soft tissue sample and a normal soft tissue sample
XX PT and comparing the gene expression, also useful in treating soft tissue
XX PT sarcoma.
XX PS Example 2; SEQ ID NO 3344; 210pp; English.
XX CC The invention relates to a novel method for detecting soft tissue sarcoma
XX CC which comprises obtaining a first soft tissue sample from an individual
XX CC and a normal soft tissue sample from the same or different individual,
XX CC determining the expression of a gene in both samples and comparing the
XX CC expression of the gene in both soft tissue samples, where a higher level
XX CC of protein expression in the first soft tissue sample indicates the
XX CC presence of soft tissue sarcoma. The method of the invention has
XX CC cytostatic applications and may be useful for detecting soft tissue
XX CC sarcoma, possibly via gene therapy or vaccine production. The nucleic
XX CC acid sequences may be useful in diagnostic and screening applications.
XX CC The current sequence is that of a human soft tissue sarcoma-upregulated
XX CC protein of the invention. The current sequence is not shown within the
XX CC specification per se but was submitted in CD format by the inventor.
XX SQ Sequence 1615 AA;
Query Match 100.0%; Score 8736; DB 8; Length 1615;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1614; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MEAAPPGPPWPLLLLLLLLALCGCPAPAAAPLFFANRRDRLVDAGGVKLESTIVVS 60
DB 1 MEAAPPGPPWPLLLLLLLLALCGCPAPAAAPLFFANRRDRLVDAGGVKLESTIVVS 60
QY 61 GLEDAAAVDFQSKGAVTWDVSEBAIKQTYLNOTGAAVQNVVIGSLVSPDGLACDWVGK 120
DB 61 GLEDAAAVDFQSKGAVTWDVSEBAIKQTYLNOTGAAVQNVVIGSLVSPDGLACDWVGK 120
QY 121 KLYWTDSETRNIEVANLNGTSRKLFWODLQPRALADPAHYMYWTDWGETPRIERAG 180
DB 121 KLYWTDSETRNIEVANLNGTSRKLFWODLQPRALADPAHYMYWTDWGETPRIERAG 180
QY 181 MDGSTRKIIVDSIYWPNGLTIDLEEQKLYWADAKLSFIHRANLDGSPRQKVVGSLTTP 240
DB 181 MDGSTRKIIVDSIYWPNGLTIDLEEQKLYWADAKLSFIHRANLDGSPRQKVVGSLTTP 240
QY 241 FALTLSGDTLYWTDWQTRSIIACNKRGTGKKEILSALYSPMDIQVLQSOERQPFHTRCE 300
DB 241 FALTLSGDTLYWTDWQTRSIIACNKRGTGKKEILSALYSPMDIQVLQSOERQPFHTRCE 300
QY 301 EDNGGCSHLCLLSPSEPFYTACPTGVQLQNGRTCKAGAEVLLARRTDLRRISLDTP 360

DB 301 EDNGGCSHLCLLSPSEPFYTACPTGVQLQNGRTCKAGAEVLLARRTDLRRISLDTP 360
QY 361 DFTDIVLQVDDIRHAIADIDYDPLEGYVYVWTDDEVRAIRRAYLDGSGAQTIVNTEINDPDG 420
DB 361 DFTDIVLQVDDIRHAIADIDYDPLEGYVYVWTDDEVRAIRRAYLDGSGAQTIVNTEINDPDG 420
QY 421 IAVDWARNLYWTDGTDRIEVTRNLNGTSRKLIVSEDLDEPRAIALHPVGMGLMYWTDWGE 480
DB 421 IAVDWARNLYWTDGTDRIEVTRNLNGTSRKLIVSEDLDEPRAIALHPVGMGLMYWTDWGE 480
QY 481 NPKIECANLDGQERRVLVNASLGNLALDQSGKLYWGDATKDKTEVINVDGTKRRTL 540
DB 481 NPKIECANLDGQERRVLVNASLGNLALDQSGKLYWGDATKDKTEVINVDGTKRRTL 540
QY 541 LEDKLPHIFGFTLLGDFIYWTDWORRSIERVHKVKSARDVIIDQLPDLMLGLKAVNVAKVV 600
DB 541 LEDKLPHIFGFTLLGDFIYWTDWORRSIERVHKVKSARDVIIDQLPDLMLGLKAVNVAKVV 600
QY 601 GTNFCADRNGGCSHLCHFTHATRCGCPIGLELLSDMKTCIVPEAFIVFTSRAAIHRISL 660
DB 601 GTNFCADRNGGCSHLCHFTHATRCGCPIGLELLSDMKTCIVPEAFIVFTSRAAIHRISL 660
QY 661 ETNNNDVAIPLTGVKEASALDFDVSNHHIYWTDVSLKTIISRAFMNGSSVEHVVEGLDYP 720
DB 661 ETNNNDVAIPLTGVKEASALDFDVSNHHIYWTDVSLKTIISRAFMNGSSVEHVVEGLDYP 720
QY 721 EGMAVDWNGKLYWADTGTNRIEVARLDGQPRQLVWRDLNPRSLALDPTKGYIYWTEW 780
DB 721 EGMAVDWNGKLYWADTGTNRIEVARLDGQPRQLVWRDLNPRSLALDPTKGYIYWTEW 780
QY 781 GKKPRIIVRAFMDGTCNMTLVKUGRANDLTIDYADORLYWTDLDTNMISSNMIGQERVV 840
DB 781 GKKPRIIVRAFMDGTCNMTLVKUGRANDLTIDYADORLYWTDLDTNMISSNMIGQERVV 840
QY 841 IADLPPFPGLTQYSDIYVWTDWNLHSIERADKTSGRNRTLIOGHLDFVMDILVFHSSRQ 900
DB 841 IADLPPFPGLTQYSDIYVWTDWNLHSIERADKTSGRNRTLIOGHLDFVMDILVFHSSRQ 900
QY 901 DGLNDCMHNNGQCQOLCLAIIPGHRCCASHYTLDPSSRNCSPPTTLLFSQSAISRM 960
DB 901 DGLNDCMHNNGQCQOLCLAIIPGHRCCASHYTLDPSSRNCSPPTTLLFSQSAISRM 960
QY 961 PDDQHSPLILPLHGRNVKAIYDPLDKFTYWDGRQNTKRAKDDGTQPVLSLSQGG 1020
DB 961 PDDQHSPLILPLHGRNVKAIYDPLDKFTYWDGRQNTKRAKDDGTQPVLSLSQGG 1020
QY 1021 NPDRQPHDLSDIYSRITLFWTCEATNTINVHRLSGEAMGVVLRGDRDKPRAIVVNAERG 1080
DB 1021 NPDRQPHDLSDIYSRITLFWTCEATNTINVHRLSGEAMGVVLRGDRDKPRAIVVNAERG 1080
QY 1081 LYFTNMQDRAAKIBRAALDGTREVLFTTGLIRVALVVDNTLGLFWVDADLKRIESCD 1140
DB 1081 LYFTNMQDRAAKIBRAALDGTREVLFTTGLIRVALVVDNTLGLFWVDADLKRIESCD 1140
QY 1141 LSGANRLTLEDANIVQPLGLTILGKHYWTDROQMIEVERKTTGDKRTRIQGRVAHLTG 1200
DB 1141 LSGANRLTLEDANIVQPLGLTILGKHYWTDROQMIEVERKTTGDKRTRIQGRVAHLTG 1200
QY 1201 IHAVEEVSLEFSAHPCARDNGGCSHICIAKDGCTPRCSPVHLVLLQNLLTCGEPTPCS 1260
DB 1201 IHAVEEVSLEFSAHPCARDNGGCSHICIAKDGCTPRCSPVHLVLLQNLLTCGEPTPCS 1260
QY 1261 PDQFACATGEIDCIPGAWRCDFPECDQSDDEGCPVCSAAQFPFCARGQCQVDLRLRCDGE 1320
DB 1261 PDQFACATGEIDCIPGAWRCDFPECDQSDDEGCPVCSAAQFPFCARGQCQVDLRLRCDGE 1320
QY 1321 ADCQDRSDEADCAICLPNQFRCASGQCVLIIKQCDSPDCIDGSDMLCMCITPPSPDS 1380
DB 1321 ADCQDRSDEADCAICLPNQFRCASGQCVLIIKQCDSPDCIDGSDMLCMCITPPSPDS 1380
QY 1381 PAHSASLGPVIGITLSLIFVMGVVFCQRVVVCQRYAGANGFPFHYVSGTGPVPLNFIAIP 1440

Db 1381 PAHSAIGPVIGIILSLFVGVYFVQVQVYAGANGPFPHEVSGTDPHVLNFIAP 1440
Qy 1441 GSGHGPTGTACGKSMSSVSLMGCGVPLXDRNHVGTGASSSSSTKATLYPPIINP 1500
Db 1441 GSGHGPTGTACGKSMSSVSLMGCGVPLXDRNHVGTGASSSSSTKATLYPPIINP 1500
Qy 1501 PPSPATPSLYNMDMFYSSNIPATAPYRPIIIRGMAPPPTTFCSTDVCDSDYSASRWKAS 1560
Db 1501 PPSPATPSLYNMDMFYSSNIPATAPYRPIIIRGMAPPPTTFCSTDVCDSDYSASRWKAS 1560
Qy 1561 KYILDNLSDSDPYPPPTPHSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615
Db 1561 KYILDNLSDSDPYPPPTPHSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615

RESULT 9

ADRI17560
ID ADRI17560 standard; protein; 1615 AA.

AC ADRI17560;
XX
DT 04-NOV-2004 (first entry)
XX
DE Human high bone mass gene, wild type allele Zmax1, protein #2.
XX
KW Human; high bone mass; Zmax1; HBM; osteoporosis; chromosome 11q13.3;
KW osteopathic; LDL receptor; bone development; metabolic bone disease; SNP;
KW single nucleotide polymorphism.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 171
FT /note= "May be Val as the result of a single nucleotide
polymorphism"

XX US6780609-B1.

XX 24-AUG-2004.

XX 05-APR-2000; 2000US-00543771.

XX 13-JAN-1998; 98US-0071449P.

XX 23-OCT-1998; 98US-0105511P.

XX 13-JAN-1999; 99US-00229319.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Carulli JP, Little RD, Recker RR, Johnson ML;
XX WPI: 2004-623529/60.

XX N-PSDB; ADRI16919.

XX New high bone mass gene of chromosome 11q13.3, encoding protein useful
for treating, diagnosing, preventing, or screening for normal and
PT abnormal conditions of bone, including metabolic bone diseases, e.g.
PT osteoporosis.

XX Disclosure; Fig 6; 284pp; English.

XX The invention relates to an isolated amino acid protein sequence selected
from an amino acid sequence appearing as ADRI16922 or an amino acid
sequence comprising or consisting of the extracellular domain of
CC ADRI16922 (amino acids 23-1385). ADRI16922 is encoded by the HBM (high bone
mass) allele of the human Zmax1 gene and has sequence similarity to LDL
CC receptors. Also disclosed are nucleic acids, proteins, cloning vectors,
CC expression vectors, transformed hosts, methods of developing
CC pharmaceutical compositions, methods of identifying molecules involved in
bone development, and methods of diagnosing and treating diseases
CC involved in bone development. Specifically disclosed is the Zmax1 gene
and the high bone mass (HBM) allele on chromosome 11q13.3 encoding
CC ADRI16922. The protein is useful for treating, diagnosing, preventing, or
CC screening for normal and abnormal conditions of bone, including metabolic

CC bone diseases, e.g. osteoporosis. The present sequence is the protein
CC encoded by the wild type human Zmax1 DNA. NOTE: Two versions of this
CC protein are present in the specification, ADRI16921 (sequence listing) and
CC ADRI17560 (shown as encoded by ADRI16919 in figure 6 and the sequence
CC listing). It is not clear which is the true protein since ADRI16921
CC contains translation exceptions and ADRI17560 doesn't.

XX Sequence 1615 AA;

Query Match 100.0%; Score 8736; DB 8; Length 1615;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1614; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MEAAPPPPPPLLLLLLLLLLALCGCPAPAAASPLLLFANRRDRLVLDAGGVKLESTIVVS 60
Db 1 MEAAPPPPPPLLLLLLLLLLALCGCPAPAAASPLLLFANRRDRLVLDAGGVKLESTIVVS 60
Qy 61 GLEDAAAVDFQSKGAVYTDVSEBAIKQTYLNQTAAGVQNVVIGSLVSPGLACDWGK 120
Db 61 GLEDAAAVDFQSKGAVYTDVSEBAIKQTYLNQTAAGVQNVVIGSLVSPGLACDWGK 120
Qy 121 KLYWTDSETRNRIEVANLNGTSRKVLFWQDLDPRAIALDPAHGYMYTWDGETPRIERAG 180
Db 121 KLYWTDSETRNRIEVANLNGTSRKVLFWQDLDPRAIALDPAHGYMYTWDGETPRIERAG 180
Qy 181 MDGSTRKIIIVDSDIYWPNGLTIDLEEQKLYWADAKLSFIHRANLDGSPQKRVGSLTTP 240
Db 181 MDGSTRKIIIVDSDIYWPNGLTIDLEEQKLYWADAKLSFIHRANLDGSPQKRVGSLTTP 240
Qy 241 FALTSGDTLYTWDQTRSIIHACNKRITGGRKEILSALYSPMDIQVLSQERQPPFHTRCE 300
Db 241 FALTSGDTLYTWDQTRSIIHACNKRITGGRKEILSALYSPMDIQVLSQERQPPFHTRCE 300
Qy 301 EDNGGCHLCLLSSEPPFYCACPTGVQLQNGRTCKAGAEVLLARTRLSLDTTP 360
Db 301 EDNGGCHLCLLSSEPPFYCACPTGVQLQNGRTCKAGAEVLLARTRLSLDTTP 360
Qy 361 DFTDIVLQVDDIRHAIADIDYDPLGYYVTDDVRAIRRAYLDGSGAOTLVNTEINDPDG 420
Db 361 DFTDIVLQVDDIRHAIADIDYDPLGYYVTDDVRAIRRAYLDGSGAOTLVNTEINDPDG 420
Qy 421 IAVDVAARNLYTDTGTDRIEVTRNGTSRKILVSEDLDEPRAIALHPVGLMYTWDGE 480
Db 421 IAVDVAARNLYTDTGTDRIEVTRNGTSRKILVSEDLDEPRAIALHPVGLMYTWDGE 480
Qy 481 NPKIECANLDGQERRVLVNASLGNPGLALDQSGKLYWGDAKTDKIEVINVDGTRRTL 540
Db 481 NPKIECANLDGQERRVLVNASLGNPGLALDQSGKLYWGDAKTDKIEVINVDGTRRTL 540
Qy 541 LEDKLPHIFGFTLLGDFIYTDWQRRSIEVHKVKSARDVIIDQLPDLMLGLKAVNVAKV 600
Db 541 LEDKLPHIFGFTLLGDFIYTDWQRRSIEVHKVKSARDVIIDQLPDLMLGLKAVNVAKV 600
Qy 601 GTNPACDRNGGCSHLCFFTHATRCGCPIGLELLSDMKTCIVPEAFVFTSRAAIIHRLS 660
Db 601 GTNPACDRNGGCSHLCFFTHATRCGCPIGLELLSDMKTCIVPEAFVFTSRAAIIHRLS 660
Qy 661 ETNNNDVAIPLTGVKEASALDFVSNNHIIYTDVSLKTIISAPFMNGSSVEHVEFGLDYP 720
Db 661 ETNNNDVAIPLTGVKEASALDFVSNNHIIYTDVSLKTIISAPFMNGSSVEHVEFGLDYP 720
Qy 721 EGMADVAMGNKLYWADTGTNRIEVARLDGQPVLRDLNPRSLALDPTKGYIYTWTEW 780
Db 721 EGMADVAMGNKLYWADTGTNRIEVARLDGQPVLRDLNPRSLALDPTKGYIYTWTEW 780
Qy 781 GKGPRIVRAFMDGTCMTLVDKVGRANDLTIDYADQRLYWTDLDTNMTIENSSNMLQERVV 840
Db 781 GKGPRIVRAFMDGTCMTLVDKVGRANDLTIDYADQRLYWTDLDTNMTIENSSNMLQERVV 840
Qy 841 IADDLPHFPGLTQVSDIYYTDMNLHLSIERADKTSGRNRTLIOGHLDPMWILVPHSSRQ 900
Db 841 IADDLPHFPGLTQVSDIYYTDMNLHLSIERADKTSGRNRTLIOGHLDPMWILVPHSSRQ 900

```
QY 901 DGLNDMHNNGCGQLCLAI PGHRCGCASHYTLDPSSRNCSPPTTLLFSSQSAISRM 960
D 901 DGLNDMHNNGCGQLCLAI PGHRCGCASHYTLDPSSRNCSPPTTLLFSSQSAISRM 960
QY 961 PDQHSPLILPLHGLRNKAIIDYDPLDKFIYWDGRQNIKRAKDDGTQPFVLTSLSQG 1020
D 961 PDQHSPLILPLHGLRNKAIIDYDPLDKFIYWDGRQNIKRAKDDGTQPFVLTSLSQG 1020
QY 1021 NPDRPHDLSIDIYRTLFWTCENTINVRHLSGEAMGVLRGRDKPRAIVNNAERG 1080
D 1021 NPDRPHDLSIDIYRTLFWTCENTINVRHLSGEAMGVLRGRDKPRAIVNNAERG 1080
QY 1081 LYFTNMQRRAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLFWVDADLKRIESC 1140
D 1081 LYFTNMQRRAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLFWVDADLKRIESC 1140
QY 1141 LSGANRLTLEDANIYVQPLGLTLGLKHLWIDRQQQMIEERVEKTDGKTRIQGRVAHLTG 1200
D 1141 LSGANRLTLEDANIYVQPLGLTLGLKHLWIDRQQQMIEERVEKTDGKTRIQGRVAHLTG 1200
QY 1201 IHAVEEVSLEBFAHPCARDNGGCSHICIAKDGTPRCSCPVLHVLQNLITCGEPPTCS 1260
D 1201 IHAVEEVSLEBFAHPCARDNGGCSHICIAKDGTPRCSCPVLHVLQNLITCGEPPTCS 1260
QY 1261 PDQFACATGEIDCIPGAWKCDGFPEDDQSDDEEGCPVCSAAQFPCCARGCQVDLRLRCDGE 1320
D 1261 PDQFACATGEIDCIPGAWKCDGFPEDDQSDDEEGCPVCSAAQFPCCARGCQVDLRLRCDGE 1320
QY 1321 ADCQDRSEADCAICLNPQFRCSAGCQVLKQCCDPPDCIDGSDLMCBITKPPSDDS 1380
D 1321 ADCQDRSEADCAICLNPQFRCSAGCQVLKQCCDPPDCIDGSDLMCBITKPPSDDS 1380
QY 1381 PAHSAIGPVGIIILSLFVGMGVYFCQVRCQYAGANGPPHRYVSGTHVPLNFTAP 1440
D 1381 PAHSAIGPVGIIILSLFVGMGVYFCQVRCQYAGANGPPHRYVSGTHVPLNFTAP 1440
QY 1441 GGSQHGPTGTIACGKSMSSVSLMGRCGVPLYDRNHTVTGASSSSSTKATLPPILNP 1500
D 1441 GGSQHGPTGTIACGKSMSSVSLMGRCGVPLYDRNHTVTGASSSSSTKATLPPILNP 1500
QY 1501 PPSPATPSLYNMDFYSSNIPATARPVYIIRGMAPPTTTCSTDVCDSDYSASRWKAS 1560
D 1501 PPSPATPSLYNMDFYSSNIPATARPVYIIRGMAPPTTTCSTDVCDSDYSASRWKAS 1560
QY 1561 KYLDLNSDSDPYPPPTPHSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615
D 1561 KYLDLNSDSDPYPPPTPHSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615

RESULT 10
ADR48211
ID ADR48211 standard; protein; 1615 AA.
XX AC
XX AC
XX AC
DT 02-DEC-2004 (first entry)
XX
DE Human high bone mass gene, wild type allele Zmax1, protein #2.
XX
KW Human; high bone mass; Zmax1; HBM; bone modulation;
KW bone development disorder; osteoporosis; chromosome 11q13.3; SNP;
KW single nucleotide polymorphism; gene therapy.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 171
FT /note= "May be Val as the result of a single nucleotide
FT polymorphism"
XX
FN US2004176582-A1.
XX
PD 09-SEP-2004.
```

```
XX 10-DEC-2003; 2003US-00731739.
PF
XX
PR 13-JAN-1998; 98US-0071449P.
PR 23-OCT-1998; 98US-0105511P.
PR 13-JAN-1999; 99US-0022931P.
PR 05-APR-2000; 2000US-00544398.
XX (GENO-) GENOME THERAPEUTICS CORP.
XX (UYCR-) UNIV CREIGHTON.
XX Carulli JP, Little RD, Recker RR, Johnson ML;
XX WPI; 2004-661408/64.
XX N-PFDB; ADR47570.
XX
XX New nucleic acid sequence encoding high bone mass, useful in diagnosing,
XX treating and/or preventing osteoporosis.
XX
XX Disclosure; Fig 6; 303pp; English.
XX
XX The invention relates to an isolated nucleic acid sequence encoding a
XX high bone mass protein (HBM). The gene exists in two alleles, Zmax1, the
XX notional wild-type (the cDNA for which appears as ADR47570 encoding
XX ADR47572) and the HBM allele (the cDNA for which appears as ADR47571
XX encoding ADR47573). The two alleles differ by a single nucleotide
XX polymorphism (G to T at position 582 of ADR47570) causing a Gly to Val
XX change at position 171 of the protein. Also included are a replicative
XX cloning vector comprising HBM/Zmax1 (and a replicon operative in an
XX isolated host cell), an expression vector comprising HBM/Zmax1 operably
XX linked to a transcription regulatory region, an isolated host cell
XX transformed with the vector(s), a method for testing a substance as a
XX therapeutic agent for bone modulation in a host, a method of identifying
XX a molecule involved in bone modulation, a method for identifying a
XX (candidate) protein involved in bone modulation, a method of testing for
XX HBM activity, a method of developing a pharmaceutical for the treatment
XX of bone development disorders, a method for treating a bone development
XX disorder in an animal, a method of altering bone development in a host, a
XX method for diagnostic screening for a genetic predisposition to a bone
XX development disorder, a diagnostic assay for bone development disorders,
XX a method of expressing the HBM protein in bone tissue, a bacterial
XX artificial chromosome comprising HBM/Zmax1 sequence (appearing as
XX ADR47574-ADR47580), a method for amplifying a nucleotide polymorphism in
XX the Zmax1 or HBM gene, a method for identifying a regulatory element of a
XX HBM gene and an isolated nucleic acid segment of at least 15 contiguous
XX nucleotides including a polymorphic site from HBM/Zmax1. The nucleic acid
XX molecule and the encoded polypeptide, composition, and methods are useful
XX in diagnosing, treating and preventing a bone development disorder, i.e.
XX osteoporosis. The gene for HBM/Zmax1 is located on chromosome 11q13.3.
XX The present sequence represents an HBM protein and encoded by the Zmax1
XX allele. NOTE: ADR47570 and ADR47571 are both stated as being 5120bp in
XX length but are not, neither do they encode accurately ADR47572 and
XX ADR47573. Both cDNAs do accurately encode 2 further versions of HBM
XX proteins appearing as ADR48211 and ADR48212.
XX
XX Sequence 1615 AA;
```

```
Query Match 100.0%; Score 8736; DB 8; Length 1615;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1614; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MEAAPGPPPLLLLLLLLLLALCCGAPAAASPLLLFANRRDVLVAGGVKLESTIVS 60
D 1 MEAAPGPPPLLLLLLLLLLALCCGAPAAASPLLLFANRRDVLVAGGVKLESTIVS 60
QY 61 GLEDAARVDQFSGAVYVTDVSEAIKQYLNQTGAQVNVVIGLVSPDGLACDWGK 120
D 61 GLEDAARVDQFSGAVYVTDVSEAIKQYLNQTGAQVNVVIGLVSPDGLACDWGK 120
QY 121 KLYWTDSETNRIEVANLNGTSRKVLFWQDLDQPPRAIADLPAHGYMYTWDGTPRIERAG 180
D 121 KLYWTDSETNRIEVANLNGTSRKVLFWQDLDQPPRAIADLPAHGYMYTWDGTPRIERAG 180
```

QY 181 MDGSTRKIIIVDSIYWPNGLTIDLEEQKLYWADAKLSPIHRANLDGSPROKQVVGSLTHP 240
DB 181 MDGSTRKIIIVDSIYWPNGLTIDLEEQKLYWADAKLSPIHRANLDGSPROKQVVGSLTHP 240
QY 241 FALTLSGDTLWTDWQTRSIIHACNKRGTGGKKEILSALYSMDIQVLSQERQPPFHTRCE 300
DB 241 FALTLSGDTLWTDWQTRSIIHACNKRGTGGKKEILSALYSMDIQVLSQERQPPFHTRCE 300
QY 301 EDNGCSHLCILSPSEPYTCACPTGVOLQNGRTCKAGAEVLLALARTDLRLISLDT 360
DB 301 EDNGCSHLCILSPSEPYTCACPTGVOLQNGRTCKAGAEVLLALARTDLRLISLDT 360
QY 361 DFTDIVLQVDDIRHAIAIDYDPLEGYVYVTDDEVRARIRAYLDGSGAQTLYNTEINDPDG 420
DB 361 DFTDIVLQVDDIRHAIAIDYDPLEGYVYVTDDEVRARIRAYLDGSGAQTLYNTEINDPDG 420
QY 421 IAVDWARNLYWTDGTDRIRIEVTRNLGTSRKILVSEDLDEPRAIALHPVWGLMTWTDWGE 480
DB 421 IAVDWARNLYWTDGTDRIRIEVTRNLGTSRKILVSEDLDEPRAIALHPVWGLMTWTDWGE 480
QY 481 NPKIECANLDQERRVLVNASLGWPNGLALDLQEGKLYWDAKTDKIEVINVDGTKRRTL 540
DB 481 NPKIECANLDQERRVLVNASLGWPNGLALDLQEGKLYWDAKTDKIEVINVDGTKRRTL 540
QY 541 LEDKLPHIFGFTLLGDFIYWTDMQRRSIRVHKVKASRDVIIDQLPDLMLKAVNAKV 600
DB 541 LEDKLPHIFGFTLLGDFIYWTDMQRRSIRVHKVKASRDVIIDQLPDLMLKAVNAKV 600
QY 601 GTNPCADRNGCCHLCFTTPHATRCGCPIGLELLSDMKTCIVPEAFVFTSRAAIIHRISL 660
DB 601 GTNPCADRNGCCHLCFTTPHATRCGCPIGLELLSDMKTCIVPEAFVFTSRAAIIHRISL 660
QY 661 ETNNNDVAIPUTGVKEASALDPVSNHHIYVTDVSLKTTISRAFWNGSSVEHVFGLDYP 720
DB 661 ETNNNDVAIPUTGVKEASALDPVSNHHIYVTDVSLKTTISRAFWNGSSVEHVFGLDYP 720
QY 721 EGMADVMMGNKLYWADTGTNRIEVARLDGQFQVILWRDLNPRSLALDPTKGIYWTW 780
DB 721 EGMADVMMGNKLYWADTGTNRIEVARLDGQFQVILWRDLNPRSLALDPTKGIYWTW 780
QY 781 GKGPRIVRAFMVGNMTCMTLVKVGANDLTIDYADQRLYVTDLTNMISSNMLGQERVV 840
DB 781 GKGPRIVRAFMVGNMTCMTLVKVGANDLTIDYADQRLYVTDLTNMISSNMLGQERVV 840
QY 841 IADLPHFPFGITQYSDIYVTDNHLSTERADKTSGRNRTLIQGHLPVMDILVPHSSRQ 900
DB 841 IADLPHFPFGITQYSDIYVTDNHLSTERADKTSGRNRTLIQGHLPVMDILVPHSSRQ 900
QY 901 DGLNDCMHNNGCCQQLCIAIPGHRGCGASHYTLDPSSRNCSPPTTFLFQKSAISRMI 960
DB 901 DGLNDCMHNNGCCQQLCIAIPGHRGCGASHYTLDPSSRNCSPPTTFLFQKSAISRMI 960
QY 961 PDDQHSPLIILPLHGLRNVAIDYDPLDKFIYVWDGRONIKRAKDDGTQPPVLTSLSGQG 1020
DB 961 PDDQHSPLIILPLHGLRNVAIDYDPLDKFIYVWDGRONIKRAKDDGTQPPVLTSLSGQG 1020
QY 1021 NPDRQPHDLSIDIVSRTLFTCEATNTINVRHLSGEMGVVLGRDRKPRAIIVNAERGY 1080
DB 1021 NPDRQPHDLSIDIVSRTLFTCEATNTINVRHLSGEMGVVLGRDRKPRAIIVNAERGY 1080
QY 1081 LYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLKFWVDADLKRIESCD 1140
DB 1081 LYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLKFWVDADLKRIESCD 1140
QY 1141 LSGANRLTLEDAIVQPIGLTILGKHLVWIDRQQMIERVKTTGDKRTRIQGRVAHLTG 1200
DB 1141 LSGANRLTLEDAIVQPIGLTILGKHLVWIDRQQMIERVKTTGDKRTRIQGRVAHLTG 1200
QY 1201 IHAVEEVSLEBSFAHPCARDNGGCSHICAKDGTGPRCSPVHLVQLNLTTCGEPPTCS 1260
DB 1201 IHAVEEVSLEBSFAHPCARDNGGCSHICAKDGTGPRCSPVHLVQLNLTTCGEPPTCS 1260
QY 1261 PDQFACATGEIDCIPGAWRCDFPECDQSDDEGCPVCSAAQFFPCARQCQVDLRLRCDGE 1320

DB 1261 PDQFACATGEIDCIPGAWRCDFPECDQSDDEGCPVCSAAQFFPCARQCQVDLRLRCDGE 1320
QY 1321 ADCQDRSDEADCAICLPNQPRCAGQCCLIKQCCDSPDCIDGSDLMCBITRPPSDDS 1380
DB 1321 ADCQDRSDEADCAICLPNQPRCAGQCCLIKQCCDSPDCIDGSDLMCBITRPPSDDS 1380
QY 1381 PAHSSAIGPVIIGIILSLFVMGVYVFCORVVCQRYAGANGPPHEYVSGTTHVPLNFIAP 1440
DB 1381 PAHSSAIGPVIIGIILSLFVMGVYVFCORVVCQRYAGANGPPHEYVSGTTHVPLNFIAP 1440
QY 1441 GGSQHGPFPTGIACGKSMMSVSLMGGRGVGLYDRNHVVTGASSSSSSSTKATLYPPIILNP 1500
DB 1441 GGSQHGPFPTGIACGKSMMSVSLMGGRGVGLYDRNHVVTGASSSSSSSTKATLYPPIILNP 1500
QY 1501 PPSPATDPSLYNMDMFYSSNIPATARPYPVYIIRGMAPPPTPCSTDVCDSDYSASRWKAS 1560
DB 1501 PPSPATDPSLYNMDMFYSSNIPATARPYPVYIIRGMAPPPTPCSTDVCDSDYSASRWKAS 1560
QY 1561 KYILDLNLSDDSDPYPPPTPHSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615
DB 1561 KYILDLNLSDDSDPYPPPTPHSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615

RESULT 11

ADR73482
ID ADR73482 standard; protein, 1615 AA.
XX ADR73482;
AC ADR73482;
XX
DT 02-DEC-2004 (first entry)
XX
DE Human low density lipoprotein receptor-related protein 5, LRP5, protein.
XX
KW LRP5; apoptosis; inhibitor; tumour; cytostatic; cancer.
XX
OS Homo sapiens.
XX
PN WO2004076682-A2.
XX
PD 10-SEP-2004.
XX
PF 26-FEB-2004; 2004WO-US0006020.
XX
PR 26-FEB-2003; 2003US-0450886P.
XX
FA (SURR-) SURROMED INC.
XX
PI Axenovich SA, Stull R, Gelman M, Chui K, Ng D;
XX
DR WPI; 2004-653428/63.
DR N-FSDB; ADR73481.
XX
PT Identifying a compound for inducing apoptosis, useful for diagnosing and
PT treating cancer, comprises identifying an inhibitor of a target.
XX
PS Claim 1; SEQ ID NO 50; 255pp; English.
XX

CC The invention relates to a novel method for identifying a compound for
CC inducing apoptosis comprising identifying an inhibitor of a target.
CC Examples of an inhibitor of a target include: angio-associated, migratory
CC cell protein (AAMP), disintegrin and metalloproteinase domain 8 (ADAM8),
CC disintegrin-like and metalloprotease (reporlysin type) with
CC thrombospondin type 1 motif, 17 (ADAMTS17), adenylate cyclase 3 (ADCY3),
CC adrenergic beta receptor kinase 1 (ADRBK1), bladder cancer associated
CC protein (BLCAP), chromosome 22 open reading frame 5 (C22orf5), CD81
CC antigen (target of antiproliferative antibody 1 (CD81), CD9 antigen (p24
CC (CD9), claudin 4 (CLDN4), chloride intracellular channel 1 (CLIC1),
CC collagen, type VI, alpha 2 (COL6A2), CTL2, endothelin converting enzyme 1
CC (ECE1), ephrinB1 (EPNBI), flotillin 2 (FLOT2), intercellular adhesion
CC molecule 3 (ICAM3), iduronate 2-sulfatase (Hunter syndrome) (IDS), jagged
CC 2 (JAG2), junctional adhesion molecule 1 (JAM1), lectin, galactoside-
CC binding soluble 3 binding protein (LGALS3BP), similar to possible G-

protein receptor (LOC146330), CGI-78 protein (LOC51107), lipoprotein lipase (LPL), low density lipoprotein receptor-related protein 5 (LRP5), Lutheran blood group (Aubergier b antigen included) (LU), membrane component, chromosomell, surface marker 1 (MLSI), serum constituent protein (MSE5), neuropathy target esterase (NTE), Homo sapiens CDNA FL31043 fig, clone HSYRA2000248 (PLEXIN A1) or Homo sapiens CDNA FLJ44113 fig, clone TESTI4046487, highly similar to Mus musculus plexin A1 (PLXNA1), protein tyrosine phosphatase, receptor type, f polypeptide (PTPRF), interacting protein (liprin), alpha 3 (PPIA3), Homo sapiens peptide-histidine transporter 4 (PTR4), mRNA (PTR4), solute carrier family 16 (monocarboxylic acid transporters) member 3 (SLC16A3), solute carrier family 1 (neutral amino acid transporter) member 5 (SLC1A5), solute carrier family 39 (zinc transporter) member 3 (SLC39A1), serine protease inhibitor, Kunitz type 2 (SPINT2), stanniocalcin 2 (STC2), tumour necrosis receptor superfamily member 21 (TNFRSF21), tumour rejection antigen (gp96) 1 (TRAI), or transient receptor potential cation channel, subfamily M member 4 (TRPM4), respectively comprising any of the 40 sequences of 87-1615 amino acids (the even numbers selected from SEQ ID:2-80), shown in the specification. The invention further comprises a method for inducing apoptosis and a method for diagnosing a tumour. The novel compounds have cytostatic activity. The method is useful for identifying a compound for inducing apoptosis. The methods and compounds are useful for diagnosing and treating cancer, and in determining the prognosis for cancer in the patient or the susceptibility of the patient to a therapeutic treatment. This sequence represents one of the 40 protein targets of the invention.

CC protein receptor (LOC146330), CGI-78 protein (LOC51107), lipoprotein
 CC lipase (LPL), low density lipoprotein receptor-related protein 5 (LRP5),
 CC Lutheran blood group (Aubergier b antigen included) (LU), membrane
 CC component, chromosomell, surface marker 1 (MLSI), serum constituent
 CC protein (MSE5), neuropathy target esterase (NTE), Homo sapiens CDNA
 CC FL31043 fig, clone HSYRA2000248 (PLEXIN A1) or Homo sapiens CDNA
 CC fig, clone TESTI4046487, highly similar to Mus musculus plexin A1
 CC (PLXNA1), protein tyrosine phosphatase, receptor type, f polypeptide
 CC (PTPRF), interacting protein (liprin), alpha 3 (PPIA3), Homo sapiens
 CC peptide-histidine transporter 4 (PTR4), mRNA (PTR4), solute carrier
 CC family 16 (monocarboxylic acid transporters) member 3 (SLC16A3), solute
 CC carrier family 1 (neutral amino acid transporter) member 5 (SLC1A5),
 CC solute carrier family 39 (zinc transporter) member 3 (SLC39A1), serine
 CC protease inhibitor, Kunitz type 2 (SPINT2), stanniocalcin 2 (STC2),
 CC tumour necrosis receptor superfamily member 21 (TNFRSF21), tumour
 CC rejection antigen (gp96) 1 (TRAI), or transient receptor potential cation
 CC channel, subfamily M member 4 (TRPM4), respectively comprising any of the
 CC 40 sequences of 87-1615 amino acids (the even numbers selected from SEQ
 CC ID:2-80), shown in the specification. The invention further comprises a
 CC method for inducing apoptosis and a method for diagnosing a tumour. The
 CC novel compounds have cytostatic activity. The method is useful for
 CC identifying a compound for inducing apoptosis. The methods and compounds
 CC are useful for diagnosing and treating cancer, and in determining the
 CC prognosis for cancer in the patient or the susceptibility of the patient
 CC to a therapeutic treatment. This sequence represents one of the 40
 CC protein targets of the invention.

XX
 SQ Sequence 1615 AA;

Query Match 100.0%; Score 8736; DB 8; Length 1615;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1614; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MEAAPGGPWPPLLLLLLLLLLALCGCPAPAAAPLFFANRRDRLVLDAGGVKLESTIVVS 60
 DB 1 MEAAPGGPWPPLLLLLLLLLLALCGCPAPAAAPLFFANRRDRLVLDAGGVKLESTIVVS 60

QY 61 GLEDAADVDFQSGAVYWTVDSEBAIKQTYLNOTGAQVNVVLSGLVSPDGLACDWGK 120
 DB 61 GLEDAADVDFQSGAVYWTVDSEBAIKQTYLNOTGAQVNVVLSGLVSPDGLACDWGK 120

QY 121 KLYTDSSTNRIEVANLNGTSRKVLFWQDLQPRALDPAHGYMYTWDGTPRIERAG 180
 DB 121 KLYTDSSTNRIEVANLNGTSRKVLFWQDLQPRALDPAHGYMYTWDGTPRIERAG 180

QY 181 MDGSTRKLIIVSDIYWPNGLIDLEEQKLYWADAKLSFIHRANLDGSPRQKVRGSLTHP 240
 DB 181 MDGSTRKLIIVSDIYWPNGLIDLEEQKLYWADAKLSFIHRANLDGSPRQKVRGSLTHP 240

QY 241 FALTLSGDTLYWTDQWTSIHACNKRKTGKKEILSALYSPMDIQVLSQERQPPFHTRCE 300
 DB 241 FALTLSGDTLYWTDQWTSIHACNKRKTGKKEILSALYSPMDIQVLSQERQPPFHTRCE 300

QY 301 EDNGCCHLCLSPSEPPYTCACPTGVQDNGRTCKAGAEVILLARRTDLRLISLTP 360
 DB 301 EDNGCCHLCLSPSEPPYTCACPTGVQDNGRTCKAGAEVILLARRTDLRLISLTP 360

QY 361 DFTDVLQVDDIRHAIDYDPLEGYVYWTDEVRIRAVYLDGSGAQTIVNTEINDPDG 420
 DB 361 DFTDVLQVDDIRHAIDYDPLEGYVYWTDEVRIRAVYLDGSGAQTIVNTEINDPDG 420

QY 421 IAVDWARNLYWTDGTDRIEVTRLNGTSRKILYSEDLDEPRALAHFVGMGLMYTWDGGE 480
 DB 421 IAVDWARNLYWTDGTDRIEVTRLNGTSRKILYSEDLDEPRALAHFVGMGLMYTWDGGE 480

QY 481 NPKIECANLDCQERRVLVNASIGWPNGLALDQEGKLYWDAKTDKIEVINVDGTKRRTL 540
 DB 481 NPKIECANLDCQERRVLVNASIGWPNGLALDQEGKLYWDAKTDKIEVINVDGTKRRTL 540

QY 541 LEDKLPPIFGFTLLGDFTLYTWDQRRSIERVHKVKASRDVLIIDQLPDLMLGLKAVNVAKV 600
 DB 541 LEDKLPPIFGFTLLGDFTLYTWDQRRSIERVHKVKASRDVLIIDQLPDLMLGLKAVNVAKV 600

RESULT 12
 AAE21741
 ID AAE21741 standard; protein; 1615 AA.
 XX

AAE21741;
 16-JUL-2002 (first entry)
 Human BSMR protein mutant, A1330L.
 Human; bone strength and mineralisation regulatory protein; BSMR;
 bone strength; mineralisation; ophthalmological; antidiabetic;
 bone density regulating transmembrane receptor; prosthetic device;
 surgical implant; diabetic retinopathy; hypertensive retinopathy;
 therapy; osteoporosis; prematurity; ocular vessel; eye disorder;
 osteopathic; mutant; mutin.
 Homo sapiens.
 OS Synthetic.
 XX
 XX Key Location/Qualifiers
 FH Misc-difference 1330
 FT /note= "Wild-type Ala substituted with Leu"
 XX WO200216553-A2.
 XX 28-FEB-2002.
 XX 17-AUG-2001; 2001WO-US041788.
 XX 18-AUG-2000; 2000US-0226119P.
 XX 22-SEP-2000; 2000US-0234337P.
 XX 13-JUL-2001; 2001US-0304851P.
 XX (AVET) AVENTIS PHARMA SA.
 XX (HARD) HARVARD COLLEGE.
 XX (UYCA-) UNIV CASE WESTERN RESERVE.
 XX Warman ML, Gong Y, Olsen BR, Rawadi G, Roman-Roman S;
 WPI; 2002-329694/36.
 XX Polynucleotide encoding bone strength and mineralization regulatory
 PT protein useful for diagnosis or therapy of osteoporosis.
 XX Disclosure; Page; 124pp; English.
 XX The invention relates to bone strength and mineralisation regulatory
 CC protein (BSMR) and its corresponding nucleic acid sequence. BSMR DNA is
 CC useful for the diagnosis or therapy of osteoporosis and for regulating
 CC (increasing) bone strength and mineralisation in a human subject by
 CC activating a bone density regulating transmembrane receptor (BSMR
 CC protein). An expression vector comprising a promoter that is operably
 CC linked to BSMR DNA is useful for modulating bone density and for
 CC enhancing bone strength and mineralisation in a mammal cell. Composition
 CC comprising a BSMR effector is useful for treating osteoporosis and is
 CC useful particularly as a coating for prosthetic devices and surgical
 CC implants. BSMR is useful for screening lead pharmaceutical agents as BSMR
 CC effectors, which may be used to treat a range of eye disorders such as
 CC diabetic retinopathy, hypertensive retinopathy and retinopathy of
 CC prematurity, in which normal vascular growth and integrity of ocular
 CC vessels is disrupted. The present sequence is human BSMR protein mutant.
 CC Note: This sequence is not shown in the specification, however it is
 CC constructed based on the human BSMR protein shown in fig 3 of the
 CC specification (AAE21730)
 XX Sequence 1615 AA;
 SQ
 Query Match 99.9%; Score 8735; DB 5; Length 1615;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1614; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MEAAPGPPWLLLLLLLLLALCGCPAPAAASPLLFANRRDRLVDAGGKLESTIVS 60
 DB 1 MEAAPGPPWLLLLLLLLLALCGCPAPAAASPLLFANRRDRLVDAGGKLESTIVS 60
 QY 61 GLEDAAAVDFQSKGAVVTVDSSEAIKQTYLNTQGAQVNVISGLVSPDGLACDWGK 120

DB 61 GLEDAAAVDFQSKGAVVTVDSSEAIKQTYLNTQGAQVNVISGLVSPDGLACDWGK 120
 QY 121 KLYWTDSETRNRIEVANLNGTSRKVLFWODLDQPAIALDPAHGMYWTDWGETPRIERAG 180
 DB 121 KLYWTDSETRNRIEVANLNGTSRKVLFWODLDQPAIALDPAHGMYWTDWGETPRIERAG 180
 QY 181 MDGSTRKLIIVDSDIYWPNGLTIDLEOKLYWADAKLSFIHRANLDGSPQKQVVEGSLTHP 240
 DB 181 MDGSTRKLIIVDSDIYWPNGLTIDLEOKLYWADAKLSFIHRANLDGSPQKQVVEGSLTHP 240
 QY 241 FALTSGDTLYWTDQTRSIIHACNKRITGGKKEILSALYSMDIQVLQSOERPPFHTRCE 300
 DB 241 FALTSGDTLYWTDQTRSIIHACNKRITGGKKEILSALYSMDIQVLQSOERPPFHTRCE 300
 QY 301 EDNGGCSHLCLLSPSEPFYTCACPTGVOLQNGRTCKAGAEVLLAARTDLRLISLTP 360
 DB 301 EDNGGCSHLCLLSPSEPFYTCACPTGVOLQNGRTCKAGAEVLLAARTDLRLISLTP 360
 QY 361 DFTDIVLQVDDIRHAIADIDPLEGYVYVTDDEVRARRAYLDGSGAQTLYNTEINDPDG 420
 DB 361 DFTDIVLQVDDIRHAIADIDPLEGYVYVTDDEVRARRAYLDGSGAQTLYNTEINDPDG 420
 QY 421 IAVDWARNLYWTDGTDRIEVTRNGTSRKILVSIEDLDEPRALALHPVGMGLMYTWDGE 480
 DB 421 IAVDWARNLYWTDGTDRIEVTRNGTSRKILVSIEDLDEPRALALHPVGMGLMYTWDGE 480
 QY 481 NPKIECANLQGERRLVNASLGNLALDQSGKLYWGDATKDKIEVINVDGTKRRTL 540
 DB 481 NPKIECANLQGERRLVNASLGNLALDQSGKLYWGDATKDKIEVINVDGTKRRTL 540
 QY 541 LEDKLPHIFGFTLLAGDFIYTDWQRRSIRVHKVAKSRDVIIDOLPDLMLGKAVNAVKV 600
 DB 541 LEDKLPHIFGFTLLAGDFIYTDWQRRSIRVHKVAKSRDVIIDOLPDLMLGKAVNAVKV 600
 QY 601 GTNFCADRNGCSHLCEFTPHATRCGCPIGLELLSDMKTCIVPEAFVFTSRAAIHRISL 660
 DB 601 GTNFCADRNGCSHLCEFTPHATRCGCPIGLELLSDMKTCIVPEAFVFTSRAAIHRISL 660
 QY 661 ETNNNDVAIPLTGVKEASALDFVSNHHIYWDVSLTKTISRPFNGSSVEHVVEGLDYP 720
 DB 661 ETNNNDVAIPLTGVKEASALDFVSNHHIYWDVSLTKTISRPFNGSSVEHVVEGLDYP 720
 QY 721 EGMADVMMGNLYWADGTNRIEVARLDGQPROVLVWRDLNPRSLALDPTKGYIYWTW 780
 DB 721 EGMADVMMGNLYWADGTNRIEVARLDGQPROVLVWRDLNPRSLALDPTKGYIYWTW 780
 QY 781 GGGKPRIVRAFMDGTNCMTLVDKVGRANDLTIDYADQRLYWTDLDTNMISSNMLQGERVV 840
 DB 781 GGGKPRIVRAFMDGTNCMTLVDKVGRANDLTIDYADQRLYWTDLDTNMISSNMLQGERVV 840
 QY 841 IADDLPHPPGLTQVSDIYIYTDWNLHLSIERADKTSGRNRTLIQGHLPDMDILVPHSSRQ 900
 DB 841 IADDLPHPPGLTQVSDIYIYTDWNLHLSIERADKTSGRNRTLIQGHLPDMDILVPHSSRQ 900
 QY 901 DGLNDCHNNGCCGOLCLAIPEGHRCGCASHYTLDPSSNCSPPTTLLFSQKSAISRWI 960
 DB 901 DGLNDCHNNGCCGOLCLAIPEGHRCGCASHYTLDPSSNCSPPTTLLFSQKSAISRWI 960
 QY 961 PDDQHSPLDILPLHGLNRVKAIDYDPLDKFKFYVWDGQNIKRAKDDGTQPPFVLTSLSQGQ 1020
 DB 961 PDDQHSPLDILPLHGLNRVKAIDYDPLDKFKFYVWDGQNIKRAKDDGTQPPFVLTSLSQGQ 1020
 QY 1021 NPDQPHDLSDIYISRTLFWTCEATNTINVHRLSGEAMGVVLRGRDRKPRAIVNAERG 1080
 DB 1021 NPDQPHDLSDIYISRTLFWTCEATNTINVHRLSGEAMGVVLRGRDRKPRAIVNAERG 1080
 QY 1081 LYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLFWVDADLKEISCD 1140
 DB 1081 LYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLFWVDADLKEISCD 1140
 QY 1141 LSGANRLTLEDANIVQPLGLTILGKHYLIDRQOQMIEVERKTTGDKRTRIQGRVAHLTG 1200

Db 1141 LSGANRLTEDANIVQPLGLTILGKLYWIDRQQQMIERVEKTTGDKRTRIQRVAHLTG 1200
 QY 1201 IHAVEVLEBFAHPCARDNGGCSHICIAKGDGTPRCSCPVHLVLLQNLITCGEPPTCS 1260
 Db 1201 IHAVEVLEBFAHPCARDNGGCSHICIAKGDGTPRCSCPVHLVLLQNLITCGEPPTCS 1260
 QY 1261 PDQFACATGEIDCIPGAWRCDFPCCDDQSBEGCPVCSAAQFPCCARGCQVDLRLRCDDGE 1320
 Db 1261 PDQFACATGEIDCIPGAWRCDFPCCDDQSBEGCPVCSAAQFPCCARGCQVDLRLRCDDGE 1320
 QY 1321 ADCQDRSDEADCAICLENQFRCSAGQCVLIKQCCDSPFCIDGSDLMCEITKPPSDDS 1380
 Db 1321 ADCQDRSDEADCAICLENQFRCSAGQCVLIKQCCDSPFCIDGSDLMCEITKPPSDDS 1380
 QY 1381 PAHSSAIGFVGIILSLFVMGVFVFCQVVCQRYAGANGPFPHEVYSGTGHVPLNFTAP 1440
 Db 1381 PAHSSAIGFVGIILSLFVMGVFVFCQVVCQRYAGANGPFPHEVYSGTGHVPLNFTAP 1440
 QY 1441 GSGQHGPTGTIACGKSMSSVSLMGGRGVPLYDRNHVYTGASSSSSSTKATLYPPILNP 1500
 Db 1441 GSGQHGPTGTIACGKSMSSVSLMGGRGVPLYDRNHVYTGASSSSSSTKATLYPPILNP 1500
 QY 1501 PPSPATPSLYNMDMFYSSNIPATAPRYPIIRGMAPPPTTTCSTDVCDSDYSASRWKAS 1560
 Db 1501 PPSPATPSLYNMDMFYSSNIPATAPRYPIIRGMAPPPTTTCSTDVCDSDYSASRWKAS 1560
 QY 1561 KYILDNSDPPYPPPTPHSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615
 Db 1561 KYILDNSDPPYPPPTPHSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615
 RESULT 13
 ADRI17561
 ID ADRI17561 standard; protein; 1615 AA.
 AC ADRI17561;
 XX
 DT 04-NOV-2004 (first entry)
 DE Human high bone mass gene, HBM allele, protein #2.
 KW Human; high bone mass; Zmax1; HBM; osteoporosis; chromosome 11q13.3;
 KW osteopathic; LDL receptor; bone development; metabolic bone disease; SNP;
 KW single nucleotide polymorphism.
 XX
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT Misc-difference 171
 FT /note= "May be Gly as the result of a single nucleotide
 FT polymorphism"
 XX
 FN US6780609-B1.
 PD 24-AUG-2004.
 XX
 PF 05-APR-2000; 2000US-00543771.
 XX
 PR 13-JAN-1998; 98US-0071449P.
 PR 23-OCT-1998; 98US-0105511P.
 PR 13-JAN-1999; 99US-00229319.
 XX
 PA (GENO-) GENOME THERAPEUTICS CORP.
 XX
 PI Carulli JP, Little RD, Recker RR, Johnson ML;
 XX
 DR WPI; 2004-623529/60.
 DR N-PSDB; ADRI16920.
 XX
 PT New high bone mass gene of chromosome 1.1Q13.3, encoding protein useful
 PT for treating, diagnosing, preventing, or screening for normal and
 PT abnormal conditions of bone, including metabolic bone diseases, e.g.
 PT osteoporosis.

XX PS Disclosure; Col 99-110; 284pp; English.
 CC The invention relates to an isolated amino acid protein sequence selected
 CC from an amino acid sequence appearing as ADRI16922 or an amino acid
 CC sequence comprising or consisting of the extracellular domain of
 CC ADRI16922(amino acids 23-1385). ADRI16922 is encoded by the HBM (high bone
 CC mass) allele of the human Zmax1 gene and has sequence similarity to LDL
 CC receptors. Also disclosed are nucleic acids, proteins, cloning vectors,
 CC expression vectors, transformed hosts, methods of developing
 CC pharmaceutical compositions, methods of identifying molecules involved in
 CC bone development, and methods of diagnosing and treating diseases
 CC involved in bone development. Specifically disclosed is the Zmax1 gene
 CC and the high bone mass (HBM) allele on chromosome 11q13.3 encoding
 CC ADRI16922. The protein is useful for treating, diagnosing, preventing, or
 CC screening for normal and abnormal conditions of bone, including metabolic
 CC bone diseases, e.g. osteoporosis. The present sequence is the protein
 CC encoded by the disease causing allele of human Zmax1. HBM. NOTE: Two
 CC versions of this protein are present in the specification, ADRI16922
 CC (sequence listing) and ADRI17561 (shown as encoded by ADRI16920 in the
 CC sequence listing). It is not clear which is the true protein since
 CC ADRI16922 contains translation exceptions and ADRI17561 doesn't.
 XX SQ Sequence 1615 AA;
 Query Match 99.9%; Score 8727; DB 8; Length 1615;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1613; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MEAAAPPGPPLLLLLLLLLLALCGCPAPAAASPLLLFANRRDVRVLVADGGVKLESTIVVS 60
 Db 1 MEAAAPPGPPLLLLLLLLLLALCGCPAPAAASPLLLFANRRDVRVLVADGGVKLESTIVVS 60
 QY 61 GLEDAAAVDFOFSKGVAVTVDVSEAIKQTYLNTGAAVQNVVLSGLVSPGLACDWGK 120
 Db 61 GLEDAAAVDFOFSKGVAVTVDVSEAIKQTYLNTGAAVQNVVLSGLVSPGLACDWGK 120
 QY 121 KLYWTDSETNRIEVANLNGTSRKVLFWQDLQOPRAIALDPAHGYMYTWDWCTETPRIERAG 180
 Db 121 KLYWTDSETNRIEVANLNGTSRKVLFWQDLQOPRAIALDPAHGYMYTWDWCTETPRIERAG 180
 QY 181 MDGSTRKIIIVDSDIYWPNGLTIDLEEQKLYWADAKLSFIHRANLDGSGFRQKVGESLTHP 240
 Db 181 MDGSTRKIIIVDSDIYWPNGLTIDLEEQKLYWADAKLSFIHRANLDGSGFRQKVGESLTHP 240
 QY 241 FALTLSGDTLYWTDWQTRSIIHACNKTGGKKEILSALYSPMDIQVLSQERQPFHTRCE 300
 Db 241 FALTLSGDTLYWTDWQTRSIIHACNKTGGKKEILSALYSPMDIQVLSQERQPFHTRCE 300
 QY 301 EDNGCSHLCLLSPSEPFYTCACPTGVQLODNGRTCKAGAEVLLARTRDLRISLDTP 360
 Db 301 EDNGCSHLCLLSPSEPFYTCACPTGVQLODNGRTCKAGAEVLLARTRDLRISLDTP 360
 QY 361 DFTDIVLQVDDIRHAIADIDYDPLEGYVYVTDDEVRAIRRAYLDGSGAQTLVNTINDPDG 420
 Db 361 DFTDIVLQVDDIRHAIADIDYDPLEGYVYVTDDEVRAIRRAYLDGSGAQTLVNTINDPDG 420
 QY 421 IAVDWARNLYWTDGTDRIEVTRNGTSRKILVSEDLDEPRALHPVGMGLMYTWDGE 480
 Db 421 IAVDWARNLYWTDGTDRIEVTRNGTSRKILVSEDLDEPRALHPVGMGLMYTWDGE 480
 QY 481 NPKIECANLDGQERRVLVNASLGNPLALDQSGKLYWGDAKTDKIEVINVDGTPKRTLL 540
 Db 481 NPKIECANLDGQERRVLVNASLGNPLALDQSGKLYWGDAKTDKIEVINVDGTPKRTLL 540
 QY 541 LEDKLPHIFGFTLLGDFIYWTWQRRSIRVHKVKSARDVIIDQLPDLMLGLKANVAKV 600
 Db 541 LEDKLPHIFGFTLLGDFIYWTWQRRSIRVHKVKSARDVIIDQLPDLMLGLKANVAKV 600
 QY 601 GTNFCADBNCGCSHLCPFTTHATECGCPIGLELLSDMKTCLVPEAFVFTSRAAHRISL 660
 Db 601 GTNFCADBNCGCSHLCPFTTHATECGCPIGLELLSDMKTCLVPEAFVFTSRAAHRISL 660

661 ETNNNDVAIPITGVKEASALDFVSNHHIYTDVSLKTIISRAFMNGSSVEHVFEGLDYP 720
661 ETNNNDVAIPITGVKEASALDFVSNHHIYTDVSLKTIISRAFMNGSSVEHVFEGLDYP 720
721 EGMADVMMKGLYADTGTNRIEVARLDGQFQVILVWRDLNPRSLALDPTKGIYTWTEW 780
721 EGMADVMMKGLYADTGTNRIEVARLDGQFQVILVWRDLNPRSLALDPTKGIYTWTEW 780
781 GKGPRIVRAFMDGTNCMTLVKVGGRANDLTIDYADQRLYTDLTNMIESNMLGQERVV 840
781 GKGPRIVRAFMDGTNCMTLVKVGGRANDLTIDYADQRLYTDLTNMIESNMLGQERVV 840
841 IADLLPFPFGLTQYSDIYIYTDWNLHLSIERADKTSGRNRTLIQGHLDVFMVMDILVFHSSRQ 900
841 IADLLPFPFGLTQYSDIYIYTDWNLHLSIERADKTSGRNRTLIQGHLDVFMVMDILVFHSSRQ 900
901 DGLNDCHMNGCCGOLCLAIPEGHRCGCASHYTLDPSSRNCSPPTFTLLFSQKSAISMI 960
901 DGLNDCHMNGCCGOLCLAIPEGHRCGCASHYTLDPSSRNCSPPTFTLLFSQKSAISMI 960
961 PDDQHSPLIILPLHGLRNVKAIYDPLDKFYIYWDGRQNIKRAKDDGTQPPFVLTSLSQG 1020
961 PDDQHSPLIILPLHGLRNVKAIYDPLDKFYIYWDGRQNIKRAKDDGTQPPFVLTSLSQG 1020
1021 NPDQPHDLSIDIYRTLFTWCEATNTINVRHLSGEAMGVVLRGDRDKPRAIVVNAERG 1080
1021 NPDQPHDLSIDIYRTLFTWCEATNTINVRHLSGEAMGVVLRGDRDKPRAIVVNAERG 1080
1081 LYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVALVVDNLTKLFWVDADLKRISCD 1140
1081 LYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVALVVDNLTKLFWVDADLKRISCD 1140
1141 LSGANRLTLEDANIVQPLGLTILGKHLWIDRQOQMIERVEKTTGDKKTRIQGRVAHUTG 1200
1141 LSGANRLTLEDANIVQPLGLTILGKHLWIDRQOQMIERVEKTTGDKKTRIQGRVAHUTG 1200
1201 IHAVEEVSLEBFAHPCARDNGGCSHICIAKGDGTPRCSPVHLVLQNLTKCGEPPTCS 1260
1201 IHAVEEVSLEBFAHPCARDNGGCSHICIAKGDGTPRCSPVHLVLQNLTKCGEPPTCS 1260
1261 PDQFACATGETDCIPGAWRCDFPECCDDSDDEBGCPCVCSAAQFCARGCQVDLRLCDGE 1320
1261 PDQFACATGETDCIPGAWRCDFPECCDDSDDEBGCPCVCSAAQFCARGCQVDLRLCDGE 1320
1321 ADCQDRSEADCAICLPNQFRCSAGQCLIKQCCDSDPDCIDGSDLMCEITKPPSDDS 1380
1321 ADCQDRSEADCAICLPNQFRCSAGQCLIKQCCDSDPDCIDGSDLMCEITKPPSDDS 1380
1381 PAHSSAIGPVIGIILSLFVMGVYFVQVRVVCQRYAGANGPFPPEHYVSGTPEHPLNFIAP 1440
1381 PAHSSAIGPVIGIILSLFVMGVYFVQVRVVCQRYAGANGPFPPEHYVSGTPEHPLNFIAP 1440
1441 GGSQHGPTGTACGKSMWSSVSLMGRCGVPLDNRHVTGASSSSSTKATLYPPLINP 1500
1441 GGSQHGPTGTACGKSMWSSVSLMGRCGVPLDNRHVTGASSSSSTKATLYPPLINP 1500
1501 PPSPATDPSLVNMDYFSSNIPATARPVPIIRGMAPPPTTTCSTDVCDSDYSASRWKAS 1560
1501 PPSPATDPSLVNMDYFSSNIPATARPVPIIRGMAPPPTTTCSTDVCDSDYSASRWKAS 1560
1561 KYLDLNSDSDPYPPPTPHSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615
1561 KYLDLNSDSDPYPPPTPHSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615

RESULT 14

ADR48212

ID ADR48212 standard; protein; 1615 AA.

XX AC

XX ADR48212;

XX 02-DEC-2004 (first entry)

XX XX

DE XX Human high bone mass gene, HBM allele, protein #2.
KW XX Human; high bone mass; Zmax1; HBM; bone modulation;
KW XX bone development disorder; osteoporosis; chromosome 11q13.3; SNP;
XX XX single nucleotide polymorphism; gene therapy.
OS OS Homo sapiens.
XX XX
FH XX Key Location/Qualifiers
FT XX Misc-difference 171
FT XX /note= "May be Gly as the result of a single nucleotide
XX XX polymorphism"
XX XX US2004176582-A1.
XX XX 09-SEP-2004.
XX XX 10-DEC-2003; 2003US-007311739.
XX XX 13-JAN-1998; 98US-0071449P.
XX XX 23-OCT-1998; 98US-0105511P.
XX XX 13-JAN-1999; 99US-00229319.
XX XX 05-APR-2000; 2000US-00544398.
XX XX (GENO-) GENOME THERAPEUTICS CORP.
XX XX (UYCR-) UNIV CREIGHTON.
XX XX Carulli JP, Little RD, Recker RR, Johnson ML;
XX XX WPI; 2004-661408/64.
XX XX N-PSDB; ADR47571.
XX XX New nucleic acid sequence encoding high bone mass, useful in diagnosing,
XX XX treating and/or preventing osteoporosis.
XX XX Disclosure; Col 99-110; 303pp; English.
XX XX The invention relates to an isolated nucleic acid sequence encoding a
XX XX high bone mass protein (HBM). The gene exists in two alleles, Zmax1, the
XX XX notional wild-type (the cDNA for which appears as ADR47570 encoding
XX XX ADR47572) and the HBM allele (the cDNA for which appears as ADR47571
XX XX encoding ADR47573). The two alleles differ by a single nucleotide
XX XX polymorphism (G to T at position 582 of ADR47570) causing a Gly to Val
XX XX change at position 171 of the protein. Also included are a replicative
XX XX cloning vector comprising HBM/Zmax1 (and a replicon operative in an
XX XX isolated host cell), an expression vector comprising HBM/Zmax1 operably
XX XX linked to a transcription regulatory region, an isolated host cell
XX XX transformed with the vector(s), a method for testing a substance as a
XX XX therapeutic agent for bone modulation in a host, a method of identifying a
XX XX molecule involved in bone modulation, a method for identifying a
XX XX (candidate) protein involved in bone modulation, a method of testing for
XX XX HBM activity, a method of developing a pharmaceutical for the treatment
XX XX of bone development disorders, a method for treating a bone development
XX XX disorder in an animal, a method of altering bone development in a host, a
XX XX method for diagnostic screening for a genetic predisposition to a bone
XX XX development disorder, a diagnostic assay for bone development disorders,
XX XX a method of expressing the HBM protein in bone tissue, a bacterial
XX XX artificial chromosome comprising HBM/Zmax1 sequence (appearing as
XX XX ADR47574-ADR47580), a method for amplifying a nucleotide polymorphism in
XX XX the Zmax1 or HBM gene, a method for identifying a regulatory element of a
XX XX HBM gene and an isolated nucleic acid segment of at least 15 contiguous
XX XX nucleotides including a polymorphic site from HBM/Zmax1. The nucleic acid
XX XX molecule and the encoded polypeptide, composition, and methods are useful
XX XX in diagnosing, treating and preventing a bone development disorder, i.e.
XX XX osteoporosis. The gene for HBM/Zmax1 is located on chromosome 11q13.3.
XX XX The present sequence represents an HBM protein and encoded by the HBM
XX XX allele. NOTE: ADR47570 and ADR47571 are both stated as being 5120bp in
XX XX length but are not, neither do they encode accurately ADR47572 and
XX XX ADR47573. Both cDNAs do accurately encode 2 further versions of HBM
XX XX proteins appearing as ADR48211 and ADR48212.
XX XX Sequence 1615 AA;

Query Match		99.9%;	Score 8727;	DB 8;	Length 1615;		
Best Local Similarity		99.9%;	Pred. No. 0;				
Matches 1613;		Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;		
QY	1	MEAAPGPPWPLLLLLLLLLLALCGCPAPAAASPALLFANRRDVRIVLDAGGVKLESTIVVS	60				
Db	1	MEAAPGPPWPLLLLLLLLLLALCGCPAPAAASPALLFANRRDVRIVLDAGGVKLESTIVVS	60				
QY	61	GLEDAADVDFQSKGAVYWDVSEAIKQTYLNOTGAAVQNVVISGLVSPDGLACDWGK	120				
Db	61	GLEDAADVDFQSKGAVYWDVSEAIKQTYLNOTGAAVQNVVISGLVSPDGLACDWGK	120				
QY	121	KLYWTDSETRNIEVANLNGTSKVLFWQDLOPRAIALDPAGHYMYWTDWGETPRIERAG	180				
Db	121	KLYWTDSETRNIEVANLNGTSKVLFWQDLOPRAIALDPAGHYMYWTDWGETPRIERAG	180				
QY	181	MDGSTRKIIVSDIYWPNGLTIDLEEQKLYWADAKLSFIHRANLDGSRQKVVESGLTTP	240				
Db	181	MDGSTRKIIVSDIYWPNGLTIDLEEQKLYWADAKLSFIHRANLDGSRQKVVESGLTTP	240				
QY	241	FALTLSGDTLWTDWQTSIHACNKRITGGRKEILSALYSPMDIQVLQOERQPPFHTRCE	300				
Db	241	FALTLSGDTLWTDWQTSIHACNKRITGGRKEILSALYSPMDIQVLQOERQPPFHTRCE	300				
QY	301	EDNGGCSHLCLLSPSEPPYTCACPTGVOLQDNGRTCKAGAEVLLARRTDLRLISLDT	360				
Db	301	EDNGGCSHLCLLSPSEPPYTCACPTGVOLQDNGRTCKAGAEVLLARRTDLRLISLDT	360				
QY	361	DFTDIVLQVDDIRHAIADYDPLEGYVYWDDEVRAIRRAYLDGSGAQTIVNTEINDPDG	420				
Db	361	DFTDIVLQVDDIRHAIADYDPLEGYVYWDDEVRAIRRAYLDGSGAQTIVNTEINDPDG	420				
QY	421	IADVARNLWTDGTDRIEVTRNGTSRKLIVSEDLDEPRATALHPVWGLMYTDWGE	480				
Db	421	IADVARNLWTDGTDRIEVTRNGTSRKLIVSEDLDEPRATALHPVWGLMYTDWGE	480				
QY	481	NPKIECANLDQERRVLVNASLWPNGLALDQEGKLYWDAKTDKIEVINVDGTKRRTL	540				
Db	481	NPKIECANLDQERRVLVNASLWPNGLALDQEGKLYWDAKTDKIEVINVDGTKRRTL	540				
QY	541	LEDKLPHIFGTLGDFIYWTDMQRRSIRVHKVKAASRDVIIDQLPDLMLGKAVNAVAV	600				
Db	541	LEDKLPHIFGTLGDFIYWTDMQRRSIRVHKVKAASRDVIIDQLPDLMLGKAVNAVAV	600				
QY	601	GTNPCADRNGCGSHLCFTPHATRCGCPIGLELLSDMKTCIVPEAFVFTSRAIHRISL	660				
Db	601	GTNPCADRNGCGSHLCFTPHATRCGCPIGLELLSDMKTCIVPEAFVFTSRAIHRISL	660				
QY	661	ETNNNDVAIPLTGVKEASALDFVSNHHIYWTDVSLKTIISRAFMNGSSVEHVVEFGLDYP	720				
Db	661	ETNNNDVAIPLTGVKEASALDFVSNHHIYWTDVSLKTIISRAFMNGSSVEHVVEFGLDYP	720				
QY	721	EGMAVDMWGNKLYWADTGTNRNIEVARLDGQFQVLRVDRDLNPRSLALDPTKGYIYTW	780				
Db	721	EGMAVDMWGNKLYWADTGTNRNIEVARLDGQFQVLRVDRDLNPRSLALDPTKGYIYTW	780				
QY	781	GGKPRIVAFWDTGNCMTLVKVGANDLTIDYADQRLYWTDLDTNMTIENSNMLGQERV	840				
Db	781	GGKPRIVAFWDTGNCMTLVKVGANDLTIDYADQRLYWTDLDTNMTIENSNMLGQERV	840				
QY	841	IADDLPHFPGLTQYSDIYWTDMNLHSIERADKTSGRNRTLIQGHLDVMDILVFHSRQ	900				
Db	841	IADDLPHFPGLTQYSDIYWTDMNLHSIERADKTSGRNRTLIQGHLDVMDILVFHSRQ	900				
QY	901	DGLNDGMHNGCGGOLCLAIIPGHRCCGASHYTLDPSSRNCSPPTTFLFQSKAISRM	960				
Db	901	DGLNDGMHNGCGGOLCLAIIPGHRCCGASHYTLDPSSRNCSPPTTFLFQSKAISRM	960				
QY	961	PDDQHSPLIPLHGLRNKALDYDPLDKFYWVDGGRONIKRAKDDGTQPVFLTSLSQG	1020				
Db	961	PDDQHSPLIPLHGLRNKALDYDPLDKFYWVDGGRONIKRAKDDGTQPVFLTSLSQG	1020				
QY	1021	NPDQPHDLSIDIYSRITFWTCEATNTINVHLSGEAMGVVLRGDRDPRAIVVNAERG	1080				

Db	1021	NPDQPHDLSIDIYSRITFWTCEATNTINVHLSGEAMGVVLRGDRDPRAIVVNAERG	1080				
QY	1081	LYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLFWVDADLKRIESC	1140				
Db	1081	LYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLFWVDADLKRIESC	1140				
QY	1141	LSGANRLTLEDANIVOPGLTILGKHLWYIDROOQMIERVEKTTGDKRTRIOGRAHUTG	1200				
Db	1141	LSGANRLTLEDANIVOPGLTILGKHLWYIDROOQMIERVEKTTGDKRTRIOGRAHUTG	1200				
QY	1201	IHAVEEVSLEEFSAHPCARDNGGCSHICIAKGDGTPRCSVHLVLLQNLITCGEPPTCS	1260				
Db	1201	IHAVEEVSLEEFSAHPCARDNGGCSHICIAKGDGTPRCSVHLVLLQNLITCGEPPTCS	1260				
QY	1261	PDQACATGEIDCIPGAWRCDFPECDQSDDEEGCPVCSAQFFPCARQCVDLRLRCDGE	1320				
Db	1261	PDQACATGEIDCIPGAWRCDFPECDQSDDEEGCPVCSAQFFPCARQCVDLRLRCDGE	1320				
QY	1321	ADCDRSDAEDCDACILPNQPRCAGGCVLTKQCCDSFPDCIDGSDLMCEITKPPSDDS	1380				
Db	1321	ADCDRSDAEDCDACILPNQPRCAGGCVLTKQCCDSFPDCIDGSDLMCEITKPPSDDS	1380				
QY	1381	PAHSSAIGPVIGIILSLFWMGVYFVQVVCQRYAGANGPPHEYYSGTTPHVLNFIAP	1440				
Db	1381	PAHSSAIGPVIGIILSLFWMGVYFVQVVCQRYAGANGPPHEYYSGTTPHVLNFIAP	1440				
QY	1441	GGSHQGFPTGIACGKSMSSVSLMGGRGVPLYDRNHVTGASSSSSSSTKATLYPPIINP	1500				
Db	1441	GGSHQGFPTGIACGKSMSSVSLMGGRGVPLYDRNHVTGASSSSSSSTKATLYPPIINP	1500				
QY	1501	PPSPATPSLYNMDMFYSSNIPATARYPRVYIIRGMAPPPTPCSTDYCDSDYSASRWKAS	1560				
Db	1501	PPSPATPSLYNMDMFYSSNIPATARYPRVYIIRGMAPPPTPCSTDYCDSDYSASRWKAS	1560				
QY	1561	KYYLDLMSDSDPYPPPTPHSQYLSAEDSCPPSPATERSYFHLFPPPPSPCTDSS	1615				
Db	1561	KYYLDLMSDSDPYPPPTPHSQYLSAEDSCPPSPATERSYFHLFPPPPSPCTDSS	1615				

RESULT 15
ABM85665
ID ABM85665 standard; protein; 1627 AA.
XX
AC ABM85665;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human protein sequence hCP1690976.
XX
KW Cytostatic; carcinoma; lymphoma; cancer; human.
XX
OS Homo sapiens.
XX
PN WO2003073826-A2.
XX
PD 12-SEP-2003.
XX
PF 28-FEB-2003; 2003WO-US006235.
XX
PR 01-MAR-2002; 2002US-00087192.
XX
PA (SAGR-) SAGRES DISCOVERY.
XX
PI Morris DW;
XX
DR WPI; 2003-328604/31.
XX
PT Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
PT comprises a nucleotide sequence.
XX
PS Claim 5; SEQ ID NO 1410; Opp; English.
XX

CC The present invention relates to novel DNA and protein sequences which
 CC are associated with carcinomas. The sequences are useful for: (i) for
 CC screening drug candidates; (ii) for screening of bioactive agent capable
 CC of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
 CC a bioactive agent capable of modulating the activity of CAP; (iv) for
 CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
 CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
 CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
 CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
 CC determining Carcinoma Associated (CA) gene copy number. In addition, the
 CC CA genes are useful as DNA vaccines and the CAP are useful as markers of
 CC carcinoma including lymphoma. The present sequence is one such CAP. Note:
 CC This patent is an equivalent to basic patent US2002182586A1, for which no
 CC sequence data was published

XX Sequence 1627 AA;

Query Match 99.8%; Score 8724; DB 7; Length 1627;
 Best Local Similarity 99.3%; Pred. No. 0;
 Matches 1615; Conservative 0; Mismatches 0; Indels 12; Gaps 1;

Qy	1	MEAAPPGPPWPLLLLLLLALCCGCPAPAAASPLLFANRRDVRVLVDAGGVKLESTIVVS	60
Db	1	MEAAPPGPPWPLLLLLLLALCCGCPAPAAASPLLFANRRDVRVLVDAGGVKLESTIVVS	60
Qy	61	GLEDAAAVDFQFSKAVYWTVDSEBAIKQTYLNOTGAQVQNVVIGSLVSPDGLACDWYVK	120
Db	61	GLEDAAAVDFQFSKAVYWTVDSEBAIKQTYLNOTGAQVQNVVIGSLVSPDGLACDWYVK	120
Qy	121	KLYWTDSTNRIEVANLNGTSRKVLFWQDLQOPRAIALDPDAHGYMYTWDGETPRIERAG	180
Db	121	KLYWTDSTNRIEVANLNGTSRKVLFWQDLQOPRAIALDPDAHGYMYTWDGETPRIERAG	180
Qy	181	MDGSTRKLIIVDSIYWPNGLTIDLEQKLYWADAKLSFIHRANLDGSPROKVVGSLTTP	240
Db	181	MDGSTRKLIIVDSIYWPNGLTIDLEQKLYWADAKLSFIHRANLDGSPROKVVGSLTTP	240
Qy	241	FALTLSGDTLYWTDQWTRSHACNKRITGKGRKEILSALYSPMDIQVLSQERQPPFHTRCE	300
Db	241	FALTLSGDTLYWTDQWTRSHACNKRITGKGRKEILSALYSPMDIQVLSQERQPPFHTRCE	300
Qy	301	EDNGCCHLLCSPEPYTCACPTGVOLQNGRTCKAGAEVLLARRTDLRLISLTP	360
Db	301	EDNGCCHLLCSPEPYTCACPTGVOLQNGRTCKAGAEVLLARRTDLRLISLTP	360
Qy	361	DFTDVLQVDDIRHAIAIDYDPLEGYVYWTDEVRARRAYLDGSGAOTLVNTEINDPDG	420
Db	361	DFTDVLQVDDIRHAIAIDYDPLEGYVYWTDEVRARRAYLDGSGAOTLVNTEINDPDG	420
Qy	421	IADVVAARNLYWTDGTDRIEVTRNGTSRKILVSEDLDEPRALHPVMGLMYTWDGE	480
Db	421	IADVVAARNLYWTDGTDRIEVTRNGTSRKILVSEDLDEPRALHPVMGLMYTWDGE	480
Qy	481	NPKIECANLGOERVLVNASLGNWGLALDQEGKLYWDAKTDKIEVINVDGTKRRTL	540
Db	481	NPKIECANLGOERVLVNASLGNWGLALDQEGKLYWDAKTDKIEVINVDGTKRRTL	540
Qy	541	LEDKLPHTFGTLLGDFIYWTDMQRRSIERVHKVKSADVLIIDQLPDLMLKAVNAVAVV	600
Db	541	LEDKLPHTFGTLLGDFIYWTDMQRRSIERVHKVKSADVLIIDQLPDLMLKAVNAVAVV	600
Qy	601	GTNPCADRNGGCHLCFFTPHATRCGCPGLELLSDMTKICVPEAFVFTSRAAIIHSL	660
Db	601	GTNPCADRNGGCHLCFFTPHATRCGCPGLELLSDMTKICVPEAFVFTSRAAIIHSL	660
Qy	661	ETNNNDVAIPLTGVEASALDFVSNHHIYWTDSLKTIISRAFMNGSSVEHVFGLDYP	720
Db	661	ETNNNDVAIPLTGVEASALDFVSNHHIYWTDSLKTIISRAFMNGSSVEHVFGLDYP	720
Qy	721	EGMAVDWNGKLYWADTGTNRIEVARLDGQFQVLRVWRDLNPRSLALDPTKGYIYWTEW	780
Db	721	EGMAVDWNGKLYWADTGTNRIEVARLDGQFQVLRVWRDLNPRSLALDPTKGYIYWTEW	780

Qy	781	GGKPRIVRAFMGDMTCMTLVKVGANDLTIDYDQRLYWTDLTNTMISSNMLQSERVV	840
Db	781	GGKPRIVRAFMGDMTCMTLVKVGANDLTIDYDQRLYWTDLTNTMISSNMLQSERVV	840
Qy	841	IADDLPHPFGLTQYSDYIYWTDMNLHSIERADKTSGRNRTLIQGHLPFMDILVPHSSRQ	900
Db	841	IADDLPHPFGLTQYSDYIYWTDMNLHSIERADKTSGRNRTLIQGHLPFMDILVPHSSRQ	900
Qy	901	DGLNDCHNNGCCQOLCLAIPEGHRCGCASHYTLDPSSRNCSPPTTFLLFQSKAISMI	960
Db	901	DGLNDCHNNGCCQOLCLAIPEGHRCGCASHYTLDPSSRNCSPPTTFLLFQSKAISMI	960
Qy	961	PDDQHSPLILPLHGLRNVAKAIDYDPLDKFYWYDGRQNIKRAKDDGT-----	1008
Db	961	PDDQHSPLILPLHGLRNVAKAIDYDPLDKFYWYDGRQNIKRAKDDGTQOAVLSLSPFVL	1020
Qy	1009	OPFVLTSLSQONPDQPHDLSIDIYSRTLFWTCEATNTINVHRLSGEAMGVLRGDRDK	1068
Db	1021	OPFVLTSLSQONPDQPHDLSIDIYSRTLFWTCEATNTINVHRLSGEAMGVLRGDRDK	1080
Qy	1069	PRAIVVNAERGYLEFTNMQDRAAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLKFW	1128
Db	1081	PRAIVVNAERGYLEFTNMQDRAAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLKFW	1140
Qy	1129	VDADLKRIESCDSGANRLTLEDANIYVQPLGLTILGKHLIYMDRQQMIERVEKTTGDKR	1188
Db	1141	VDADLKRIESCDSGANRLTLEDANIYVQPLGLTILGKHLIYMDRQQMIERVEKTTGDKR	1200
Qy	1189	TRIOGRAVHLTGIIHAEVSEVLEEFSAHPCARDNGGCSHICIAKGGTTPRCSPVHLVLLQ	1248
Db	1201	TRIOGRAVHLTGIIHAEVSEVLEEFSAHPCARDNGGCSHICIAKGGTTPRCSPVHLVLLQ	1260
Qy	1249	NLLTCGGPPTCPQOFACATGEIDCIPGAWRCDFPECCDDQDEEGCPVCSAAQPPCARG	1308
Db	1261	NLLTCGGPPTCPQOFACATGEIDCIPGAWRCDFPECCDDQDEEGCPVCSAAQPPCARG	1320
Qy	1309	QCVDLRLRCDEADCDQRSDEADCAICLPNQRCASQCQVLIKQCCDSPDCIDGSDCL	1368
Db	1321	QCVDLRLRCDEADCDQRSDEADCAICLPNQRCASQCQVLIKQCCDSPDCIDGSDCL	1380
Qy	1369	MCEITKPPSDSPAHSSAIGPVGIIILSLFWGCVYVQVQVQRYAGANGPPPHETYS	1428
Db	1381	MCEITKPPSDSPAHSSAIGPVGIIILSLFWGCVYVQVQVQRYAGANGPPPHETYS	1440
Qy	1429	GTPHVPLNFIAPGSGQHPFTGACGKSMSSVSLMGGRGVPLYDRNHVTGASSSSSS	1488
Db	1441	GTPHVPLNFIAPGSGQHPFTGACGKSMSSVSLMGGRGVPLYDRNHVTGASSSSSS	1500
Qy	1489	TKATLYPPIILNPPSPATDPSLYNMDFYSSNIPATAPYRPIYIRGMAPPTTTCSTDVC	1548
Db	1501	TKATLYPPIILNPPSPATDPSLYNMDFYSSNIPATAPYRPIYIRGMAPPTTTCSTDVC	1560
Qy	1549	DSYASASRWKASKYLLDNLSDSDPYPPPTPHSQYLSAEDSCPPSPATERSYFHLFPBPP	1608
Db	1561	DSYASASRWKASKYLLDNLSDSDPYPPPTPHSQYLSAEDSCPPSPATERSYFHLFPBPP	1620
Qy	1609	SPCTDSS 1615	
Db	1621	SPCTDSS 1627	

Search completed: February 17, 2005, 01:31:02
 Job time : 210 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 17, 2005, 01:16:38 ; Search time 224 Seconds
(without alignments)
3691.999 Million cell updates/sec

Title: US-09-931-375A-2

Perfect score: 8740

Sequence: 1 MEAAPGPPWPLLLLLLLLL.....TERSYFLPPPPSPCTDSS 1615

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8740	100.0	1615	2 QUES7	Ques7 homo sapien
2	8736	100.0	1615	1 LRP5 HUMAN	Q75197 homo sapien
3	8259.5	94.5	1614	1 LRP5 MOUSE	Q91V00 mus musculus
4	7132.5	81.6	1605	2 Q8AYF1	Q8AYF1 xenopus lae
5	6153.5	70.4	1613	1 LRP6 HUMAN	Q75581 homo sapien
6	6097.5	69.8	1613	1 LRP6 MOUSE	Q88572 mus musculus
7	6084	69.6	1613	2 Q8AYF0	Q8AYF0 xenopus lae
8	3527.5	40.4	1678	2 Q9V600	Q9V600 drosophila
9	3526.5	40.3	1678	2 Q9V600	Q9V600 drosophila
10	3521.5	40.3	1678	2 Q9V600	Q9V600 drosophila
11	3262.5	37.3	1698	2 Q7PV65	Q7PV65 anopheles g
12	2788	31.9	1246	2 Q6AWJ8	Q6AWJ8 drosophila
13	2687	30.7	1905	1 LRP4 RAT	Q9QYPI rattus norv
14	2687	30.7	1905	2 Q76LJ2	Q76LJ2 rattus norv
15	2686	30.7	1950	1 LRP4 HUMAN	Q75096 homo sapien
16	2671	30.6	1905	1 LRP4 MOUSE	Q8V156 mus musculus
17	2654	30.4	527	2 Q77501	Q77501 oryctolagus
18	2266.5	25.9	1768	2 Q7QEK9	Q7QEK9 anopheles g
19	2228	25.5	2009	2 Q9VXN0	Q9VXN0 drosophila
20	2193	25.1	4544	1 LRP1 HUMAN	Q07954 homo sapien
21	2180	24.9	4543	1 LRPI_CHICK	P98157 gallus gall
22	2165	24.8	4545	2 Q91Z77	Q91Z77 mus musculus
23	2165	24.8	4545	2 Q920T4	Q920T4 mus musculus
24	2113	24.2	4545	2 Q61291	Q61291 mus musculus
25	2018	23.1	1731	2 Q8WY30	Q8WY30 homo sapien
26	2008	23.0	4599	1 LR1B HUMAN	Q9NZR2 homo sapien
27	1997	22.8	4599	1 LR1B_MOUSE	Q91I18 mus musculus
28	1876.5	21.5	4655	1 LRP2_HUMAN	P98164 homo sapien
29	1876.5	21.5	4655	2 Q7Z5C0	Q7Z5C0 homo sapien
30	1872.5	21.4	4655	2 Q7Z5C1	Q7Z5C1 homo sapien
31	1856	21.2	4660	1 LRP2_RAT	P98158 rattus norv

32 1779.5 20.4 4569 2 Q7PS35
33 1762.5 20.2 4699 2 Q9V383
34 1759.5 20.1 4547 2 Q9W343
35 1708 19.5 4569 2 Q7PV66
36 1458 16.7 4753 1 LRP_CAEEL
37 1044.5 12.0 1581 2 Q73809
38 1024 11.7 252 2 Q9NSY4
39 980.5 11.2 1859 2 Q7PS28
40 969 11.1 1809 2 Q8MP02
41 914.5 10.5 1847 2 Q76952
42 909.5 10.4 1984 1 YL_DROME
43 902 10.3 1650 2 Q9QV76
44 876.5 10.0 881 2 Q8WY31
45 859 9.8 1537 2 Q8WY29

Q7ps35 anopheles g
Q9v383 drosophila
Q9w343 drosophila
Q7pv66 anopheles g
Q04833 caenorhabdi
Q73809 fugu rubrip
Q9nsy4 homo sapien
Q7ps28 anopheles g
Q8mp02 periplaneta
Q76952 aedes aegypt
P98163 drosophila
Q9qvt6 rattus sp.
Q8wy31 homo sapien
Q8wy29 homo sapien

ALIGNMENTS

RESULT 1

Q9UES7 PRELIMINARY; PRT; 1615 AA.
AC Q9UES7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE LDL receptor member LR3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9908902; PubMed=9790987; DOI=10.1006/bbrc.1998.9545;
RA Dong Y., Lathrop W., Weaver D., Qiu Q., Cini J., Bertolini D.,
RA Chen D.;
RT "Molecular cloning and characterization of LR3, a novel LDL receptor
family protein with mitogenic activity.";
RL Biochem. Biophys. Res. Commun. 251:784-790(1998).
DR EMBL; AF077820; AAC72791.1; -.
DR HSP; P98162; LURF.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR002172; LDL_receptor_A.
DR InterPro; IPR000033; Ldl_receptor_rep.
DR Pfam; PF00008; EGF; 3.
DR Pfam; PF00057; Ldl_recept_a; 3.
DR Pfam; PF00058; Ldl_recept_b; 19.
DR PRINTS; PR00261; LDLRECEPTOR.
DR SMART; SM00181; EGF; 4.
DR SMART; SM00192; LDLa; 3.
DR SMART; SM00135; LY; 20.
DR PROSITE; PS01209; LDLRA_1; 3.
DR PROSITE; PS50068; LDLRA_2; 3.
KW Receptor.
SQ SEQUENCE 1615 AA; 179143 MW; 8BA25D07F51E02CA CRC64;

Query Match 100.0%; Score 8740; DB 2; Length 1615;
Best Local Similarity 100.0%; Pred.No. 0;
Matches 1615; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEAAPGPPWPLLLLLLLLLLCCCPAPAAASPLLLFANRRDVLVADGGVKLESTIVVS 60
Db 1 MEAAPGPPWPLLLLLLLLLLCCCPAPAAASPLLLFANRRDVLVADGGVKLESTIVVS 60
Qy 61 GLEDAADVDFQSKGAVYWTVDVSEAIKQTYLNTGAAVQNVISGLVSPDGLACDWGK 120
Db 61 GLEDAADVDFQSKGAVYWTVDVSEAIKQTYLNTGAAVQNVISGLVSPDGLACDWGK 120
Qy 121 KLYWTDSETNRIEYANLNGTSRKVLPWQDLDPRAIALDPAHGYMYTWDGTPRIERAG 180
Db 121 KLYWTDSETNRIEYANLNGTSRKVLPWQDLDPRAIALDPAHGYMYTWDGTPRIERAG 180

Db 121 KLYWTDSETRIEVANLNGTSRKVLFWQDLQDRAIALDPAGHYMYWTDWGTGPRIERAG 180
QY 181 MDGSTRKIIVSDIYWPNGLTITDLBEOKLYWADAKLSFTHRANLDGSRQKQVVEGSLTHP 240
Db 181 MDGSTRKIIVSDIYWPNGLTITDLBEOKLYWADAKLSFTHRANLDGSRQKQVVEGSLTHP 240
QY 241 FALTLSGDTLYWTDWQTSRIHACNKRITGKKEIILSALYSMDIOVLQOERQPPFHTTCE 300
Db 241 FALTLSGDTLYWTDWQTSRIHACNKRITGKKEIILSALYSMDIOVLQOERQPPFHTTCE 300
QY 301 EDNGCCHLCLLSPSEPFYTCACPTGVQOLQDNGRTCKAGAEVLLARRTDLRISLDT 360
Db 301 EDNGCCHLCLLSPSEPFYTCACPTGVQOLQDNGRTCKAGAEVLLARRTDLRISLDT 360
QY 361 DFTDVLQVDDIRHAIADYDPLEGYVYVWTDDEVRAIRRAYLDGSGAOTLVNTEINDPDG 420
Db 361 DFTDVLQVDDIRHAIADYDPLEGYVYVWTDDEVRAIRRAYLDGSGAOTLVNTEINDPDG 420
QY 421 IAVDWVARNLYWTDGTDRIEVTRLNGTSRKILVSEDLDEPRATLHPVWGLMYWTDWGE 480
Db 421 IAVDWVARNLYWTDGTDRIEVTRLNGTSRKILVSEDLDEPRATLHPVWGLMYWTDWGE 480
QY 481 NPKIECANLDGQERRVLYVNASLGNPNGLALDQEGKLYWDAKTDKIEVINVDGTKRRTL 540
Db 481 NPKIECANLDGQERRVLYVNASLGNPNGLALDQEGKLYWDAKTDKIEVINVDGTKRRTL 540
QY 541 LEDKLPHFIFGFTLLGDFIYWTDWQRRSIRVHKVYKASRDVLIIDQLPDLMLGLKAVNVAKV 600
Db 541 LEDKLPHFIFGFTLLGDFIYWTDWQRRSIRVHKVYKASRDVLIIDQLPDLMLGLKAVNVAKV 600
QY 601 GTNPCADNRGCGSHLCFFTPHATRCGCPGLELLSDMKTCIVPEAFVFTSRAAHRISL 660
Db 601 GTNPCADNRGCGSHLCFFTPHATRCGCPGLELLSDMKTCIVPEAFVFTSRAAHRISL 660
QY 661 ETNNNDVAIPUTGVKEASALDPDVSNHHIYWTDVSLKTTISRAFMNGSSVERVVFGLDYP 720
Db 661 ETNNNDVAIPUTGVKEASALDPDVSNHHIYWTDVSLKTTISRAFMNGSSVERVVFGLDYP 720
QY 721 EGMADVMMGNLYWADTGTNRIEVARLDGQROVLVWDLNPSRLALDPTKGYIYVTEW 780
Db 721 EGMADVMMGNLYWADTGTNRIEVARLDGQROVLVWDLNPSRLALDPTKGYIYVTEW 780
QY 781 GKGPRIVRAFMFGDNCMTLVKVGANDLTIDYADQRLYWTDLTNMISSNMLGQERVV 840
Db 781 GKGPRIVRAFMFGDNCMTLVKVGANDLTIDYADQRLYWTDLTNMISSNMLGQERVV 840
QY 841 IADDLPHFPGLTQYSYDIYVWTDNLHLSIERADKTSGRNRTLIOGHLPFVMDILVFHSSRQ 900
Db 841 IADDLPHFPGLTQYSYDIYVWTDNLHLSIERADKTSGRNRTLIOGHLPFVMDILVFHSSRQ 900
QY 901 DGLNDCMHNNGCGCOLCIAIPGHRCCASHYTLDPGSRNCSPPTFLFSQKSAISRM 960
Db 901 DGLNDCMHNNGCGCOLCIAIPGHRCCASHYTLDPGSRNCSPPTFLFSQKSAISRM 960
QY 961 PDDQHSPLILPLHGLRNKAIIDYDPLDKFYWVDGQNIKRAKDDGTQPPVLTSLSQGG 1020
Db 961 PDDQHSPLILPLHGLRNKAIIDYDPLDKFYWVDGQNIKRAKDDGTQPPVLTSLSQGG 1020
QY 1021 NPDQPHDLSIDIYSRTLFWTCEATNTINVRHLSGEMGVVLRGDRPKPRAIVVNAERG 1080
Db 1021 NPDQPHDLSIDIYSRTLFWTCEATNTINVRHLSGEMGVVLRGDRPKPRAIVVNAERG 1080
QY 1081 LYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLFWADLAKRIESCD 1140
Db 1081 LYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLFWADLAKRIESCD 1140
QY 1141 LSGANRLTLEDANIVQPLGLTILGKHLVYIDRQOQMIERVEKTTGDKRTRIQGRVAHLTG 1200
Db 1141 LSGANRLTLEDANIVQPLGLTILGKHLVYIDRQOQMIERVEKTTGDKRTRIQGRVAHLTG 1200
QY 1201 IHAVEEVSLEEFSAHPCARDNGGCSHICIAKGDGTPRCSPVHLVLLQNLITCGEPPTCS 1260
Db 1201 IHAVEEVSLEEFSAHPCARDNGGCSHICIAKGDGTPRCSPVHLVLLQNLITCGEPPTCS 1260

QY 1261 PDQFACATGEIDCIPGAWRCDFPECDQSDDEEGCPVCSAAQFPCCARQCQVDLRLRCDGE 1320
Db 1261 PDQFACATGEIDCIPGAWRCDFPECDQSDDEEGCPVCSAAQFPCCARQCQVDLRLRCDGE 1320
QY 1321 ADCQDRSDEADCDALCPNQFRCSAGCQVLKQOCDSPDCIDGSDLMCMCITKPPSDS 1380
Db 1321 ADCQDRSDEADCDALCPNQFRCSAGCQVLKQOCDSPDCIDGSDLMCMCITKPPSDS 1380
QY 1381 PAHSSAIGPVTIGIILSLFVWGGVYFVQRVVQRYAGANGPFPHEYYSGTHVPLNFIAP 1440
Db 1381 PAHSSAIGPVTIGIILSLFVWGGVYFVQRVVQRYAGANGPFPHEYYSGTHVPLNFIAP 1440
QY 1441 GGSQHGPTGTIACGKSMSSVSLMGGRGGVPLYDRNHVHTGASSSSSTKATLYPPIILNP 1500
Db 1441 GGSQHGPTGTIACGKSMSSVSLMGGRGGVPLYDRNHVHTGASSSSSTKATLYPPIILNP 1500
QY 1501 PPSPATDPSLYNMDMYSSNIPATARPVPYIIRGMAPPPTPCSTDVDCSDYSASRWKAS 1560
Db 1501 PPSPATDPSLYNMDMYSSNIPATARPVPYIIRGMAPPPTPCSTDVDCSDYSASRWKAS 1560
QY 1561 KYLLDLNSDSDPYPPPTPHSQYLSAEDSCPPSPATERSYFHLFPPPPSPCTDSS 1615
Db 1561 KYLLDLNSDSDPYPPPTPHSQYLSAEDSCPPSPATERSYFHLFPPPPSPCTDSS 1615

RESULT 2

LRP5 HUMAN
ID LRP5 HUMAN STANDARD; PRT; 1615 AA.
AC 075197; 096TD6; Q9UP66;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Low-density lipoprotein receptor-related protein 5 precursor.
GN Name=LRP5; Synonyms=LRP7;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Osteoblast;
RX MEDLINE=98382578; PubMed=9714764; DOI=10.1016/S0378-1119(98)00311-4;
RA Hey P.J., Twells R.C.J., Metzker M.S., Nakagawa Y., Brown S.D.,
Kawaguchi Y., Cox R., Xie G., Dugan V., Hammond H., Metzker M.D.,
Todd J.A., Hess J.F.;
RT "Cloning of a novel member of the low-density lipoprotein receptor
family.";
RL Gene 216:103-111(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21295044; PubMed=11401438; DOI=10.1006/geno.2000.6492;
RA Twells R.C.J., Metzker M.L., Brown S.D., Cox R., Garey C., Hammond H.,
Hey P.J., Levy E., Nakagawa Y., Phillips M.S., Todd J.A., Hess J.F.;
RT "The sequence and gene characterization of a 400-kb candidate region
for IDDM4 on chromosome 11q13.";
RL Genomics 72:231-242(2001).
RN [3]
RP SEQUENCE FROM N.A.
RX PubMed=12509515; DOI=10.1073/pnas.013792100;
RA Fujino T., Asaba H., Kang M.J., Ikeda Y., Sone H., Takada S.,
Kim D.H., Ioka R.X., Ono M., Tomoyori H., Okubo M., Murase T.,
Kamataki A., Yamamoto J., Magoori K., Takahashi S., Miyamoto Y.,
Oishi H., Nose M., Okazaki M., Usui K., Imaizumi K., Yanagisawa M.,
Sakai J., Yamamoto T.T.;
RT "Low-density lipoprotein receptor-related protein 5 (LRP5) is
essential for normal cholesterol metabolism and glucose-induced
insulin secretion.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:229-234(2003).
RN [4]
RP FUNCTION, PHOSPHORYLATION, AND INTERACTION WITH AXIN.
RX PubMed=14731402; DOI=10.1016/S1097-2765(03)00484-2;
RA Tamai K., Zeng X., Liu C., Zhang X., Harada Y., Chang Z., He X.;

RT "A mechanism for Wnt coreceptor activation.";
RL Mol. Cell 13:149-156(2004).
RN [5]
RP REVIEW.
RX PubMed=15084453; DOI=10.1242/dev.01117;
RA He X., Semenov M., Tamai K., Zeng X.;
RT "LDL receptor-related proteins 5 and 6 in Wnt/beta-catenin signaling:
RT arrows point the way.";
RL Development 131:1663-1677(2004).
RN [6]
RP VARIANTS MET-173; HIS-1168 AND GLY-1361.
RX PubMed=15024691; DOI=10.1086/383202;
RA Toomes C., Bottomley H.M., Jackson R.M., Towns K.V., Scott S.,
RA Mackey D.A., Craig J.E., Jiang L., Yang Z., Trembath R., Woodruff G.,
RA Gregory-Evans C.Y., Gregory-Evans K., Parker M.J., Black G.C.,
RA Downey L.M., Zhang K., Inglehearn C.F.;
RT "Mutations in LRP5 or FZD4 underlie the common familial exudative
RT vitreoretinopathy locus on chromosome 11q.";
RL Am. J. Hum. Genet. 74:721-730(2004).
CC -1- FUNCTION: Involved in the Wnt/beta catenin signaling pathway,
CC probably by acting as a coreceptor together with Frizzled for Wnt.
CC -1- SUBUNIT: Interacts with different Wnt/Frizzled complexes.
CC interacts with axin.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -1- TISSUE SPECIFICITY: Widely expressed, with the highest level of
CC expression in the liver.
CC -1- PTM: Phosphorylation of the PPPSP motif creates an inducible
CC docking site for axin.
CC -1- POLYMORPHISM: The Met-17; His-1168 and Gly-1361 variants were
CC found in family with FEVER [MIM:605750]. However, asymptomatic
CC individuals were also found to have the mutations. It is thus
CC unknown whether they are disease-causing mutations.
CC -1- DISEASE: Defects in LRP5 may be a cause of familial exudative
CC vitreoretinopathy (FEVR) [MIM:605750]. FEVR is a disorder of the
CC retinal vasculature characterized by an abrupt cessation of the
CC growth of peripheral retinal capillaries. Autosomal dominant
CC inheritance is the most common mode of segregation, although
CC recessive and X-linked pedigrees are also seen. The phenotype is
CC variable. In severe cases, the formation of fibrovascular vitreous
CC membranes that cause retinal traction may result in the
CC displacement of the macula, the presence of retinal folds, and/or
CC the detachment of the retina, any of which may cause a decrease in
CC visual acuity sufficient to be labeled legal blindness. Patients
CC with mild forms of the disease are asymptomatic, and their only
CC disease-related abnormality is an arc of avascular retina in the
CC extreme temporal periphery.
CC -1- DISEASE: Defects in LRP5 may be a cause of osteoporosis
CC pseudoglioma syndrome (OPPG) [MIM:259770]. OPPG is a recessive
CC disorder characterized by very low bone mass and blindness.
CC Individually with OPPG are prone to develop bone fractures and
CC deformations and have various eye abnormalities, including
CC phthisis bulbi, retinal detachments, falciform folds or persistent
CC vitreal vasculature.
CC -1- SIMILARITY: Belongs to the LDLR family.
CC -1- SIMILARITY: Contains 4 EGF-like domains.
CC -1- SIMILARITY: Contains 3 LDL-receptor class A domains.
CC -1- SIMILARITY: Contains 19 LDL-receptor class B domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF064548; AAC36467.1; -;
DR EMBL; AF283321; AAK52433.1; -;
DR EMBL; AF283320; AAK52433.1; JOINED.
DR EMBL; AB017498; BAA33051.1; -;
DR PIN; JEO372; JEO372.
DR HSSP; P98162; 1JRF.
DR Genew; HGNC:6697; LRP5.

DR MIM; 603506; -;
DR MIM; 605750; -;
DR MIM; 259770; -;
DR GO; GO:0016021; C: integral to membrane; TAS.
DR GO; GO:0006629; P: lipid metabolism; TAS.
DR GO; GO:0008284; P: positive regulation of cell proliferation; TAS.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR008210; EGF.
DR InterPro; IPR002172; LDL_receptor_A.
DR InterPro; IPR000033; Ldl_receptor_rep.
DR InterPro; IPR011042; TolB_C.
DR Pfam; PF00008; EGF_4.
DR Pfam; PF00057; Ldl_recept_a; 3.
DR Pfam; PF00058; Ldl_recept_b; 19.
DR PRINTS; PR00261; LDLRECEPTOR.
DR SMART; SM00181; EGF; 4.
DR SMART; SM00192; LDLA; 3.
DR SMART; SM00135; LY; 20.
DR PROSITE; PS00022; EGF_1; FALSE_NEG.
DR PROSITE; PS01186; EGF_2; FALSE_NEG.
DR PROSITE; PS00026; EGF_3; FALSE_NEG.
DR PROSITE; PS01187; EGF_CA; FALSE_NEG.
DR PROSITE; PS01209; LDLRA_1; 3.
DR PROSITE; PS00068; LDLRA_2; 3.
KW Developmental protein; EGF-like domain; Endocytosis; Glycoprotein;
KW Phosphorylation; Polymorphism; Receptor; Repeat; Signal;
KW Transmembrane; Wnt signaling pathway.
FT SIGNAL 1 31
FT CHAIN 32 1615
FT DOMAIN 32 1384
FT TRANSSEM 1385 1407
FT DOMAIN 1408 1615
FT DOMAIN 75 118
FT DOMAIN 120 161
FT DOMAIN 163 205
FT DOMAIN 207 246
FT DOMAIN 248 289
FT DOMAIN 295 337
FT DOMAIN 385 426
FT DOMAIN 428 469
FT DOMAIN 471 513
FT DOMAIN 515 558
FT DOMAIN 601 641
FT DOMAIN 687 728
FT DOMAIN 730 771
FT DOMAIN 773 814
FT DOMAIN 816 853
FT DOMAIN 856 897
FT DOMAIN 902 942
FT DOMAIN 989 1034
FT DOMAIN 1036 1077
FT DOMAIN 1079 1122
FT DOMAIN 1124 1162
FT DOMAIN 1165 1207
FT DOMAIN 1213 1254
FT DOMAIN 1258 1296
FT DOMAIN 1297 1333
FT DOMAIN 1335 1371
FT DOMAIN 1495 1610
FT REPEAT 78 81
FT REPEAT 123 126
FT REPEAT 166 169
FT REPEAT 251 254
FT REPEAT 388 391
FT REPEAT 431 434
FT REPEAT 474 477
FT REPEAT 559 562
FT REPEAT 690 693
FT REPEAT 819 822
FT REPEAT 859 862
FT DOMAIN 1384 1407
FT DOMAIN 1408 1615
FT DOMAIN 75 118
FT DOMAIN 120 161
FT DOMAIN 163 205
FT DOMAIN 207 246
FT DOMAIN 248 289
FT DOMAIN 295 337
FT DOMAIN 385 426
FT DOMAIN 428 469
FT DOMAIN 471 513
FT DOMAIN 515 558
FT DOMAIN 601 641
FT DOMAIN 687 728
FT DOMAIN 730 771
FT DOMAIN 773 814
FT DOMAIN 816 853
FT DOMAIN 856 897
FT DOMAIN 902 942
FT DOMAIN 989 1034
FT DOMAIN 1036 1077
FT DOMAIN 1079 1122
FT DOMAIN 1124 1162
FT DOMAIN 1165 1207
FT DOMAIN 1213 1254
FT DOMAIN 1258 1296
FT DOMAIN 1297 1333
FT DOMAIN 1335 1371
FT DOMAIN 1495 1610
FT REPEAT 78 81
FT REPEAT 123 126
FT REPEAT 166 169
FT REPEAT 251 254
FT REPEAT 388 391
FT REPEAT 431 434
FT REPEAT 474 477
FT REPEAT 559 562
FT REPEAT 690 693
FT REPEAT 819 822
FT REPEAT 859 862
FT DOMAIN 1384 1407
FT DOMAIN 1408 1615
FT DOMAIN 75 118
FT DOMAIN 120 161
FT DOMAIN 163 205
FT DOMAIN 207 246
FT DOMAIN 248 289
FT DOMAIN 295 337
FT DOMAIN 385 426
FT DOMAIN 428 469
FT DOMAIN 471 513
FT DOMAIN 515 558
FT DOMAIN 601 641
FT DOMAIN 687 728
FT DOMAIN 730 771
FT DOMAIN 773 814
FT DOMAIN 816 853
FT DOMAIN 856 897
FT DOMAIN 902 942
FT DOMAIN 989 1034
FT DOMAIN 1036 1077
FT DOMAIN 1079 1122
FT DOMAIN 1124 1162
FT DOMAIN 1165 1207
FT DOMAIN 1213 1254
FT DOMAIN 1258 1296
FT DOMAIN 1297 1333
FT DOMAIN 1335 1371
FT DOMAIN 1495 1610
FT REPEAT 78 81
FT REPEAT 123 126
FT REPEAT 166 169
FT REPEAT 251 254
FT REPEAT 388 391
FT REPEAT 431 434
FT REPEAT 474 477
FT REPEAT 559 562
FT REPEAT 690 693
FT REPEAT 819 822
FT REPEAT 859 862
FT DOMAIN 1384 1407
FT DOMAIN 1408 1615
FT DOMAIN 75 118
FT DOMAIN 120 161
FT DOMAIN 163 205
FT DOMAIN 207 246
FT DOMAIN 248 289
FT DOMAIN 295 337
FT DOMAIN 385 426
FT DOMAIN 428 469
FT DOMAIN 471 513
FT DOMAIN 515 558
FT DOMAIN 601 641
FT DOMAIN 687 728
FT DOMAIN 730 771
FT DOMAIN 773 814
FT DOMAIN 816 853
FT DOMAIN 856 897
FT DOMAIN 902 942
FT DOMAIN 989 1034
FT DOMAIN 1036 1077
FT DOMAIN 1079 1122
FT DOMAIN 1124 1162
FT DOMAIN 1165 1207
FT DOMAIN 1213 1254
FT DOMAIN 1258 1296
FT DOMAIN 1297 1333
FT DOMAIN 1335 1371
FT DOMAIN 1495 1610
FT REPEAT 78 81
FT REPEAT 123 126
FT REPEAT 166 169
FT REPEAT 251 254
FT REPEAT 388 391
FT REPEAT 431 434
FT REPEAT 474 477
FT REPEAT 559 562
FT REPEAT 690 693
FT REPEAT 819 822
FT REPEAT 859 862
FT DOMAIN 1384 1407
FT DOMAIN 1408 1615
FT DOMAIN 75 118
FT DOMAIN 120 161
FT DOMAIN 163 205
FT DOMAIN 207 246
FT DOMAIN 248 289
FT DOMAIN 295 337
FT DOMAIN 385 426
FT DOMAIN 428 469
FT DOMAIN 471 513
FT DOMAIN 515 558
FT DOMAIN 601 641
FT DOMAIN 687 728
FT DOMAIN 730 771
FT DOMAIN 773 814
FT DOMAIN 816 853
FT DOMAIN 856 897
FT DOMAIN 902 942
FT DOMAIN 989 1034
FT DOMAIN 1036 1077
FT DOMAIN 1079 1122
FT DOMAIN 1124 1162
FT DOMAIN 1165 1207
FT DOMAIN 1213 1254
FT DOMAIN 1258 1296
FT DOMAIN 1297 1333
FT DOMAIN 1335 1371
FT DOMAIN 1495 1610
FT REPEAT 78 81
FT REPEAT 123 126
FT REPEAT 166 169
FT REPEAT 251 254
FT REPEAT 388 391
FT REPEAT 431 434
FT REPEAT 474 477
FT REPEAT 559 562
FT REPEAT 690 693
FT REPEAT 819 822
FT REPEAT 859 862
FT DOMAIN 1384 1407
FT DOMAIN 1408 1615
FT DOMAIN 75 118
FT DOMAIN 120 161
FT DOMAIN 163 205
FT DOMAIN 207 246
FT DOMAIN 248 289
FT DOMAIN 295 337
FT DOMAIN 385 426
FT DOMAIN 428 469
FT DOMAIN 471 513
FT DOMAIN 515 558
FT DOMAIN 601 641
FT DOMAIN 687 728
FT DOMAIN 730 771
FT DOMAIN 773 814
FT DOMAIN 816 853
FT DOMAIN 856 897
FT DOMAIN 902 942
FT DOMAIN 989 1034
FT DOMAIN 1036 1077
FT DOMAIN 1079 1122
FT DOMAIN 1124 1162
FT DOMAIN 1165 1207
FT DOMAIN 1213 1254
FT DOMAIN 1258 1296
FT DOMAIN 1297 1333
FT DOMAIN 1335 1371
FT DOMAIN 1495 1610
FT REPEAT 78 81
FT REPEAT 123 126
FT REPEAT 166 169
FT REPEAT 251 254
FT REPEAT 388 391
FT REPEAT 431 434
FT REPEAT 474 477
FT REPEAT 559 562
FT REPEAT 690 693
FT REPEAT 819 822
FT REPEAT 859 862
FT DOMAIN 1384 1407
FT DOMAIN 1408 1615
FT DOMAIN 75 118
FT DOMAIN 120 161
FT DOMAIN 163 205
FT DOMAIN 207 246
FT DOMAIN 248 289
FT DOMAIN 295 337
FT DOMAIN 385 426
FT DOMAIN 428 469
FT DOMAIN 471 513
FT DOMAIN 515 558
FT DOMAIN 601 641
FT DOMAIN 687 728
FT DOMAIN 730 771
FT DOMAIN 773 814
FT DOMAIN 816 853
FT DOMAIN 856 897
FT DOMAIN 902 942
FT DOMAIN 989 1034
FT DOMAIN 1036 1077
FT DOMAIN 1079 1122
FT DOMAIN 1124 1162
FT DOMAIN 1165 1207
FT DOMAIN 1213 1254
FT DOMAIN 1258 1296
FT DOMAIN 1297 1333
FT DOMAIN 1335 1371
FT DOMAIN 1495 1610
FT REPEAT 78 81
FT REPEAT 123 126
FT REPEAT 166 169
FT REPEAT 251 254
FT REPEAT 388 391
FT REPEAT 431 434
FT REPEAT 474 477
FT REPEAT 559 562
FT REPEAT 690 693
FT REPEAT 819 822
FT REPEAT 859 862
FT DOMAIN 1384 1407
FT DOMAIN 1408 1615
FT DOMAIN 75 118
FT DOMAIN 120 161
FT DOMAIN 163 205
FT DOMAIN 207 246
FT DOMAIN 248 289
FT DOMAIN 295 337
FT DOMAIN 385 426
FT DOMAIN 428 469
FT DOMAIN 471 513
FT DOMAIN 515 558
FT DOMAIN 601 641
FT DOMAIN 687 728
FT DOMAIN 730 771
FT DOMAIN 773 814
FT DOMAIN 816 853
FT DOMAIN 856 897
FT DOMAIN 902 942
FT DOMAIN 989 1034
FT DOMAIN 1036 1077
FT DOMAIN 1079 1122
FT DOMAIN 1124 1162
FT DOMAIN 1165 1207
FT DOMAIN 1213 1254
FT DOMAIN 1258 1296
FT DOMAIN 1297 1333
FT DOMAIN 1335 1371
FT DOMAIN 1495 1610
FT REPEAT 78 81
FT REPEAT 123 126
FT REPEAT 166 169
FT REPEAT 251 254
FT REPEAT 388 391
FT REPEAT 431 434
FT REPEAT 474 477
FT REPEAT 559 562
FT REPEAT 690 693
FT REPEAT 819 822
FT REPEAT 859 862
FT DOMAIN 1384 1407
FT DOMAIN 1408 1615
FT DOMAIN 75 118
FT DOMAIN 120 161
FT DOMAIN 163 205
FT DOMAIN 207 246
FT DOMAIN 248 289
FT DOMAIN 295 337
FT DOMAIN 385 426
FT DOMAIN 428 469
FT DOMAIN 471 513
FT DOMAIN 515 558
FT DOMAIN 601 641
FT DOMAIN 687 728
FT DOMAIN 730 771
FT DOMAIN 773 814
FT DOMAIN 816 853
FT DOMAIN 856 897
FT DOMAIN 902 942
FT DOMAIN 989 1034
FT DOMAIN 1036 1077
FT DOMAIN 1079 1122
FT DOMAIN 1124 1162
FT DOMAIN 1165 1207
FT DOMAIN 1213 1254
FT DOMAIN 1258 1296
FT DOMAIN 1297 1333
FT DOMAIN 1335 1371
FT DOMAIN 1495 1610
FT REPEAT 78 81
FT REPEAT 123 126
FT REPEAT 166 169
FT REPEAT 251 254
FT REPEAT 388 391
FT REPEAT 431 434
FT REPEAT 474 477
FT REPEAT 559 562
FT REPEAT 690 693
FT REPEAT 819 822
FT REPEAT 859 862
FT DOMAIN 1384 1407
FT DOMAIN 1408 1615
FT DOMAIN 75 118
FT DOMAIN 120 161
FT DOMAIN 163 205
FT DOMAIN 207 246
FT DOMAIN 248 289
FT DOMAIN 295 337
FT DOMAIN 385 426
FT DOMAIN 428 469
FT DOMAIN 471 513
FT DOMAIN 515 558
FT DOMAIN 601 641
FT DOMAIN 687 728
FT DOMAIN 730 771
FT DOMAIN 773 814
FT DOMAIN 816 853
FT DOMAIN 856 897
FT DOMAIN 902 942
FT DOMAIN 989 1034
FT DOMAIN 1036 1077
FT DOMAIN 1079 1122
FT DOMAIN 1124 1162
FT DOMAIN 1165 1207
FT DOMAIN 1213 1254
FT DOMAIN 1258 1296
FT DOMAIN 1297 1333
FT DOMAIN 1335 1371
FT DOMAIN 1495 1610
FT REPEAT 78 81
FT REPEAT 123 126
FT REPEAT 166 169
FT REPEAT 251 254
FT REPEAT 388 391
FT REPEAT 431 434
FT REPEAT 474 477
FT REPEAT 559 562
FT REPEAT 690 693
FT REPEAT 819 822
FT REPEAT 859 862
FT DOMAIN 1384 1407
FT DOMAIN 1408 1615
FT DOMAIN 75 118
FT DOMAIN 120 161
FT DOMAIN 163 205
FT DOMAIN 207 246
FT DOMAIN 248 289
FT DOMAIN 295 337
FT DOMAIN 385 426
FT DOMAIN 428 469
FT DOMAIN 471 513
FT DOMAIN 515 558
FT DOMAIN 601 641
FT DOMAIN 687 728
FT DOMAIN 730 771
FT DOMAIN 773 814
FT DOMAIN 816 853
FT DOMAIN 856 897
FT DOMAIN 902 942
FT DOMAIN 989 1034
FT DOMAIN 1036 1077
FT DOMAIN 1079 1122
FT DOMAIN 1124 1162
FT DOMAIN 1165 1207
FT DOMAIN 1213 1254
FT DOMAIN 1258 1296
FT DOMAIN 1297 1333
FT DOMAIN 1335 1371
FT DOMAIN 1495 1610
FT REPEAT 78 81
FT REPEAT 123 126
FT REPEAT 166 169
FT REPEAT 251 254
FT REPEAT 388 391
FT REPEAT 431 434
FT REPEAT 474 477
FT REPEAT 559 562
FT REPEAT 690 693
FT REPEAT 819 822
FT REPEAT 859 862
FT DOMAIN 1384 1407
FT DOMAIN 1408 1615
FT DOMAIN 75 118
FT DOMAIN 120 161
FT DOMAIN 163 205
FT DOMAIN 207 246
FT DOMAIN 248 289
FT DOMAIN 295 337
FT DOMAIN 385 426
FT DOMAIN 428 469
FT DOMAIN 471 513
FT DOMAIN 515 558
FT DOMAIN 601 641
FT DOMAIN 687 728
FT DOMAIN 730 771
FT DOMAIN 773 814
FT DOMAIN 816 853
FT DOMAIN 856 897
FT DOMAIN 902 942
FT DOMAIN 989 1034
FT DOMAIN 1036 1077
FT DOMAIN 1079 1122
FT DOMAIN 1124 1162
FT DOMAIN 1165 1207
FT DOMAIN 1213 1254
FT DOMAIN 1258 1296
FT DOMAIN 1297 1333
FT DOMAIN 1335 1371
FT DOMAIN 1495 1610
FT REPEAT 78 81
FT REPEAT 123 126
FT REPEAT 166 169
FT REPEAT 251 254
FT REPEAT 388 391
FT REPEAT 431 434
FT REPEAT 474 477
FT REPEAT 559 562
FT REPEAT 690 693
FT REPEAT 819 822
FT REPEAT 859 862
FT DOMAIN 1384 1407
FT DOMAIN 1408 1615
FT DOMAIN 75 118
FT DOMAIN 120 161
FT DOMAIN 163 205
FT DOMAIN 207 246
FT DOMAIN 248 289
FT DOMAIN 295 337
FT DOMAIN 385 426
FT DOMAIN 428 469
FT DOMAIN 471 513
FT DOMAIN 515 558
FT DOMAIN 601 641
FT DOMAIN 687 728
FT DOMAIN 730 771
FT DOMAIN 773 814
FT DOMAIN 816 853
FT DOMAIN 856 897
FT DOMAIN 902 942
FT DOMAIN 989 1034
FT DOMAIN 1036 1077
FT DOMAIN 1079 1122
FT DOMAIN 1124 1162
FT DOMAIN 1165 1207
FT DOMAIN 1213 1254
FT DOMAIN 1258 1296
FT DOMAIN 1297 1333
FT DOMAIN 1335 1371
FT DOMAIN 1495 1610
FT REPEAT 78 81
FT REPEAT 123 126
FT REPEAT 166 169
FT REPEAT 251 254
FT REPEAT 388 391
FT REPEAT 431 434
FT REPEAT 474 477
FT REPEAT 559 562
FT REPEAT 690 693
FT REPEAT 819 822
FT REPEAT 859 862
FT DOMAIN 1384 1407
FT DOMAIN 1408 1615
FT DOMAIN 75 118
FT DOMAIN 120 161
FT DOMAIN 163 205
FT DOMAIN 207 246
FT DOMAIN 248 289
FT DOMAIN 295 337
FT DOMAIN 385 426
FT DOMAIN 428 469
FT DOMAIN 471 513
FT DOMAIN 515 558
FT DOMAIN 601 641
FT DOMAIN 687 728
FT DOMAIN 730 771
FT DOMAIN 773 814
FT DOMAIN 816 853
FT DOMAIN 856 897
FT DOMAIN 902 942
FT DOMAIN 989 1034
FT DOMAIN 1036 1077
FT DOMAIN 1079 1122
FT DOMAIN 1124 1162
FT DOMAIN 1165 1207
FT DOMAIN 1213 1254
FT DOMAIN 1258 1296
FT DOMAIN 1297 1333
FT DOMAIN 1335 1371
FT DOMAIN 1495 1610
FT REPEAT 78 81
FT REPEAT 123 126
FT REPEAT 166 169
FT REPEAT 251 254
FT REPEAT 388 391
FT REPEAT 431 434
FT REPEAT 474 477
FT REPEAT 559 562
FT REPEAT 690 693
FT REPEAT 819 822
FT REPEAT 859 862
FT DOMAIN 1384 1407
FT DOMAIN 1408 1615
FT DOMAIN 75 118
FT DOMAIN 120 161
FT DOMAIN 163 205
FT DOMAIN 207 246
FT DOMAIN 248 289
FT DOMAIN 295 337
FT DOMAIN 385 426
FT DOMAIN 428 469
FT DOMAIN 471 513
FT DOMAIN 515 558
FT DOMAIN 601 641
FT DOMAIN 687 728
FT DOMAIN 730 771
FT DOMAIN 773 814
FT DOMAIN 816 853
FT DOMAIN 856 897
FT DOMAIN 902 942
FT DOMAIN 989 1034
FT DOMAIN 1036 1077
FT DOMAIN 1079 1122
FT DOMAIN 1124 1162
FT DOMAIN 1165 1207
FT DOMAIN 1213 1254
FT DOMAIN 1258 1296
FT DOMAIN 1297 1333
FT DOMAIN 1335 1371
FT DOMAIN 1495 1610
FT REPEAT 78 81
FT REPEAT 123 126
FT REPEAT 166 169
FT REPEAT 251 254
FT REPEAT 388 391
FT REPEAT 431 434
FT REPEAT 474 477
FT REPEAT 559 562
FT REPEAT 690 693
FT REPEAT 819 822
FT REPEAT 859 862
FT DOMAIN 1384 1407
FT DOMAIN 1408 1615
FT DOMAIN 75 118
FT DOMAIN 120 161
FT DOMAIN 163 205
FT DOMAIN 207 246
FT DOMAIN 248 289
FT DOMAIN 295 337
FT DOMAIN 385 426
FT DOMAIN 428 469
FT DOMAIN 471 513
FT DOMAIN 515 558
FT DOMAIN 601 641
FT DOMAIN 687 728
FT DOMAIN 730 771
FT DOMAIN 773 814
FT DOMAIN 816 853
FT DOMAIN 856 897
FT DOMAIN 902 942
FT DOMAIN 989 1034
FT DOMAIN 1036 1077
FT DOMAIN 1079 1122
FT DOMAIN 1124 1162
FT DOMAIN 1165 1207
FT DOMAIN 1213 1254
FT DOMAIN 1258 1296
FT DOMAIN 1297 1333
FT DOMAIN 1335 1371
FT DOMAIN 1495 1610
FT REPEAT 78 81
FT REPEAT 123 126
FT REPEAT 166 169
FT REPEAT 251 254
FT REPEAT 388 391
FT REPEAT 431 434
FT REPEAT 474 477
FT REPEAT 559 562
FT REPEAT 690 693
FT REPEAT 819 822
FT REPEAT 859 862
FT DOMAIN 1384 1407
FT DOMAIN 1408 1615
FT DOMAIN 75 118
FT DOMAIN 120 161
FT DOMAIN 163 205
FT DOMAIN 207 246
FT DOMAIN 248 289
FT DOMAIN 295 337
FT DOMAIN 385 426
FT DOMAIN 428 469
FT DOMAIN 471 513
FT DOMAIN 515 558
FT DOMAIN 601 641
FT DOMAIN 687 728
FT DOMAIN 730 771
FT DOMAIN 773 814
FT DOMAIN 816 853
FT DOMAIN 856 897
FT DOMAIN 902 942
FT DOMAIN 989 1034
FT DOMAIN 1036 1077
FT DOMAIN 1079 1122
FT DOMAIN 1124 1162
FT DOMAIN 1165 1207
FT DOMAIN 1213 1254
FT DOMAIN 1258 1296
FT DOMAIN 1297 1333
FT DOMAIN 1335 1371
FT DOMAIN 1495 1610
FT REPEAT 78 81
FT REPEAT 123 126
FT REPEAT 166 169
FT REPEAT 251 254
FT REPEAT 388 391
FT REPEAT 431 434
FT REPEAT 474 477
FT REPEAT 559 562
FT REPEAT 690 693
FT REPEAT 819 822
FT REPEAT 859 862
FT DOMAIN 1384 1407
FT DOMAIN 1408 1615
FT DOMAIN 75 118
FT DOMAIN 120 161
FT DOMAIN 163 205
FT DOMAIN 207 246
FT DOMAIN 248 289
FT DOMAIN 295 337
FT DOMAIN 385 426
FT DOMAIN 428 469
FT DOMAIN 471 513
FT DOMAIN 515 558
FT DOMAIN 601 641
FT DOMAIN 687 728
FT DOMAIN 730 771
FT DOMAIN 773 814
FT DOMAIN 816 853
FT DOMAIN 856 897
FT DOMAIN 902 942
FT DOMAIN 989 1034
FT DOMAIN 1036 1077
FT DOMAIN 1079 1122
FT DOMAIN 1124 1162
FT DOMAIN 1165 1207
FT DOMAIN 1213 1254
FT DOMAIN 1258 1296
FT DOMAIN 1297 1333
FT DOMAIN 1335 1371
FT DOMAIN 1495 1610
FT REPEAT 78 81
FT REPEAT 123 126
FT REPEAT 166 169
FT REPEAT 251 254
FT REPEAT 388 391
FT REPEAT 431 434
FT REPEAT 474 477
FT REPEAT 559 562
FT REPEAT 690 693
FT REPEAT 819 822
FT REPEAT 859 862
FT DOMAIN 1384 1407
FT DOMAIN 1408 1615
FT DOMAIN 75 118
FT DOMAIN 120 161
FT DOMAIN 163 205
FT DOMAIN 207 246
FT DOMAIN 248 289
FT DOMAIN 295 337
FT DOMAIN 385 426
FT DOMAIN 428 469
FT DOMAIN 471 513
FT DOMAIN 515 558
FT DOMAIN 601 641
FT DOMAIN 687 728
FT DOMAIN 730 771
FT DOMAIN 773 814
FT DOMAIN 816 853
FT DOMAIN 856 897
FT DOMAIN 902 942
FT DOMAIN 989 1034
FT DOMAIN 1036 1077
FT DOMAIN 1079 1122
FT DOMAIN 1124 1162
FT DOMAIN 1165 1207
FT DOMAIN 1213 1254
FT DOMAIN 1258 1296
FT DOMAIN 1297 1333
FT DOMAIN 1335 1371
FT DOMAIN 1495 1610
FT REPEAT 78 81
FT REPEAT 123 126
FT REPEAT 166 169
FT REPEAT 251 254
FT REPEAT 388 391
FT REPEAT 431 434
FT REPEAT 474 477
FT REPEAT 559 562
FT REPEAT 690 693
FT REPEAT 819 822
FT REPEAT 859 862
FT DOMAIN 1384 1407
FT DOMAIN 1408 1615
FT DOMAIN 75 118
FT DOMAIN 120 161
FT DOMAIN 163 205
FT DOMAIN 207 246
FT DOMAIN 248 289
FT DOMAIN 295 337
FT DOMAIN 385 426
FT DOMAIN 428 469
FT DOMAIN 471 513
FT DOMAIN 515 558
FT DOMAIN 601 641
FT DOMAIN 687 728
FT DOMAIN 730 771
FT DOMAIN 773 814
FT DOMAIN 816 853
FT DOMAIN 856 897
FT DOMAIN 902 942
FT DOMAIN 989 1034
FT DOMAIN 1036 1077
FT DOMAIN 1079 1122
FT DOMAIN 1124 1162
FT DOMAIN 1165 1207
FT DOMAIN 1213 1254
FT DOMAIN 1258 1296
FT DOMAIN 1297 1333
FT DOMAIN 1335 1371
FT DOMAIN 1495 1610
FT REPEAT 78 81
FT REPEAT 123 126
FT REPEAT 166 169
FT REPEAT 251 254
FT REPEAT 388 391
FT REPEAT 431 434
FT REPEAT 474 477
FT REPEAT 559 562
FT REPEAT 690 693
FT REPEAT 819 822
FT REPEAT

FT	SITE	1500	1504	PPSP motif.
FT	SITE	1606	1610	PPSP motif.
FT	DISULFID	239	310	By similarity.
FT	DISULFID	306	321	By similarity.
FT	DISULFID	323	336	By similarity.
FT	DISULFID	605	616	By similarity.
FT	DISULFID	612	625	By similarity.
FT	DISULFID	627	640	By similarity.
FT	DISULFID	906	917	By similarity.
Query Match				
Best Local Similarity 100.0%; Score 8736; DB 1; Length 1615;				
Matches 1614; Conservative 99.9%; Pred. No. 0;				
Mismatches 1; Indels 0; Gaps 0;				
QY	1	MEAPPPGPPWPLLLLLLLLLLALCGCPAPAPAAAPLPPANRRDRLVLDAGGVKLESTIVVS	60	
DB	1	MEAPPPGPPWPLLLLLLLLLLALCGCPAPAPAAAPLPPANRRDRLVLDAGGVKLESTIVVS	60	
QY	61	GLEDAAAVDFOFSKGVVTVDSBEAIKQTYLNOTGAAVQNVVLSGLVSPDGLACDWGK	120	
DB	61	GLEDAAAVDFOFSKGVVTVDSBEAIKQTYLNOTGAAVQNVVLSGLVSPDGLACDWGK	120	
QY	121	KLYWTDSTNRIEVANLNGTSRKVLFWQDLQPRALALDPAGHYMYWTDWGETPRIERAG	180	
DB	121	KLYWTDSTNRIEVANLNGTSRKVLFWQDLQPRALALDPAGHYMYWTDWGETPRIERAG	180	
QY	181	MDGSTRKLIIVDSIYWPNGLTIDLEOKLYWADAKLSFIHRANLDGSPRQKVVGSLTHP	240	
DB	181	MDGSTRKLIIVDSIYWPNGLTIDLEOKLYWADAKLSFIHRANLDGSPRQKVVGSLTHP	240	
QY	241	FALTLSGDTLWTDQWTRSHACNKRKTGKKEKILSALYSPMDIQVLSQERQPPFHTRCE	300	
DB	241	FALTLSGDTLWTDQWTRSHACNKRKTGKKEKILSALYSPMDIQVLSQERQPPFHTRCE	300	
QY	301	EDNGGCSHLCLSPSEPPYTCACPTGVOLQNGRTCKAGAEVLLARTRDLRISLDT	360	
DB	301	EDNGGCSHLCLSPSEPPYTCACPTGVOLQNGRTCKAGAEVLLARTRDLRISLDT	360	
QY	361	DFTDIVLQVDDIRHAIAIDYDPLSGYVYVTDDEVRATRAVLDGSGAQTLYNTEINPDG	420	
DB	361	DFTDIVLQVDDIRHAIAIDYDPLSGYVYVTDDEVRATRAVLDGSGAQTLYNTEINPDG	420	
QY	421	IADVARNLYWTDGTDRIEVTRNGTSRKILVSEDLDEPRAIALHPVGLMYWTDWGE	480	
DB	421	IADVARNLYWTDGTDRIEVTRNGTSRKILVSEDLDEPRAIALHPVGLMYWTDWGE	480	
QY	481	NPKTECANLDQERRVLVNASLGNPGLALDLOEGKLYWDAKTKIEVINVDGTKRRTL	540	
DB	481	NPKTECANLDQERRVLVNASLGNPGLALDLOEGKLYWDAKTKIEVINVDGTKRRTL	540	
QY	541	LEDKLPHPFGFTLLGDFTYWTDWRRSRIERVHKVKSARDVLIIDQLPDLMLGLKAVNVAKV	600	
DB	541	LEDKLPHPFGFTLLGDFTYWTDWRRSRIERVHKVKSARDVLIIDQLPDLMLGLKAVNVAKV	600	
QY	601	GTNPACDRNGGCSHLCPFFTHATRCGCPIGLELLSDMKTKTCVPEAFVLTFSRAIHRISL	660	
DB	601	GTNPACDRNGGCSHLCPFFTHATRCGCPIGLELLSDMKTKTCVPEAFVLTFSRAIHRISL	660	
QY	661	ETNNVDVAIPLTGVKEASALDFVSNHHIYWTDVSLKTIISRAFNWGSSEVHVVEFGLDYP	720	
DB	661	ETNNVDVAIPLTGVKEASALDFVSNHHIYWTDVSLKTIISRAFNWGSSEVHVVEFGLDYP	720	
QY	721	EGMAVDMWGNLYWADTGTNRRIEVARLDGQPROVLVWRDLNPRSLALDPTKGVIVYVTEW	780	
DB	721	EGMAVDMWGNLYWADTGTNRRIEVARLDGQPROVLVWRDLNPRSLALDPTKGVIVYVTEW	780	
QY	781	GGKPRIVAFMDGNCMTLVKVGRLNDLTDYADQRLYWTDLTDTNNIENSSNMLGQERVV	840	
DB	781	GGKPRIVAFMDGNCMTLVKVGRLNDLTDYADQRLYWTDLTDTNNIENSSNMLGQERVV	840	
QY	841	IADLPHPFGFTQVSDYIYVTDNNLHSGIERADKTSGRNRTLIOGHLDVFMVDILVFHSSRQ	900	
DB	841	IADLPHPFGFTQVSDYIYVTDNNLHSGIERADKTSGRNRTLIOGHLDVFMVDILVFHSSRQ	900	

RESULT 3
LRP5 MOUSE
ID LRP5_MOUSE STANDARD; PRT; 1614 AA.
AC Q91VNO; O88883; Q9R208;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Low-density lipoprotein receptor-related protein 5 precursor (LRP7) (Lr3).
GN Name=Lrp5; Synonyms=Lrp7;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Liver;
RX MEDLINE=98382578; PubMed=9714764; DOI=10.1016/S0378-1119(98)00311-4;
RA Hey P.J., Twells R.C.J., Phillips M.S., Nakagawa Y., Brown S.D.,
RA Kawaguchi Y., Cox R.J., Xie G., Dugan V., Hammond H., Metzker M.L.,
RA Todd J.A., Hess J.F.;
RT "Cloning of a novel member of the low-density lipoprotein receptor family.";

Gene 216:103-111(1998).

[2]

SEQUENCE FROM N.A. (ISOFORM 1), AND ALTERNATIVE SPLICING.

TISSUE=Liver;

MEDLINE=99168901; PubMed=10049586; DOI=10.1006/geno.1998.5688;

Chen D., Lathrop W., Dong Y.;

"Molecular cloning of mouse Lrp7(lr3) cDNA and chromosomal mapping of orthologous genes in mouse and human.;"

Genomic 55:314-321(1999).

[3]

SEQUENCE FROM N.A. (ISOFORM 1).

TISSUE=Breast tumor;

CC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,

Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Prange C.,

Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,

Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,

Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.B.,

Schneerch A., Schein J.E., Jones S.J.M., Marra M.A.;

"Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.;"

Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

CC -1- FUNCTION: Involved in the Wnt/beta catenin signaling pathway, Wnt probably by acting as a coreceptor together with Frizzled for Wnt (By similarity).

CC -1- SUBUNIT: Interacts with different Wnt/Frizzled complexes.

CC -1- SUBCELLULAR LOCATION: Type I membrane protein (Potential).

CC -1- ALTERNATIVE PRODUCTS:

CC Event=Alternative splicing; Named isoforms=1;

CC Comment=A number of isoforms are produced;

CC Name=1;

CC IsoId=Q91VN0-1; Sequence=Displayed;

CC -1- TISSUE SPECIFICITY: Widely expressed, with the highest expression levels in liver, heart, and lung and the lowest levels in brain and spleen.

CC -1- DEVELOPMENTAL STAGE: Expressed before or on embryonic day 7.

CC -1- PTM: Phosphorylation of the PPPSP motif creates an inducible docking site for axin (By similarity).

CC -1- SIMILARITY: Belongs to the LDLR family.

CC -1- SIMILARITY: Contains 4 EGF-like domains.

CC -1- SIMILARITY: Contains 19 LDL-receptor class A domains.

CC -1- SIMILARITY: Contains 19 LDL-receptor class B domains.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AF064984; AAC36468.1; -

DR EMBL; AF077847; AAC70183.1; -

DR EMBL; BC011374; AAH11374.1; -

DR HSSP; P01130; 1AJJ.

DR MGD; MGI:1278315; Lrp5.

DR InterPro; IPR006209; EGF like.

DR InterPro; IPR006210; IEGF.

DR InterPro; IPR002172; LDL_receptor A.

DR InterPro; IPR000033; Ldl_receptor_rep.

PFam; PF00008; EGF; 4.

DR PFam; PF00057; Ldl_recept_a; 3.

DR PFam; PF00058; Ldl_recept_b; 19.

DR PRINTS; PR00261; LDLRECEPTOR.

DR SMART; SM00181; EGF; 4.

DR SMART; SM00192; LDLa; 3.

DR SMART; SM00135; LY; 20.

DR PROSITE; PS00022; EGF_1; FALSE NEG.

DR PROSITE; PS01186; EGF_2; FALSE NEG.

DR PROSITE; PS00026; EGF_3; FALSE NEG.

DR PROSITE; PS01187; EGF_CA; FALSE NEG.

DR PROSITE; PS01209; LDLa_1; 3.

DR PROSITE; PS00068; LDLa_2; 3.

KW Alternative splicing; Developmental protein; EGF-like domain; Endocytosis; Glycoprotein; Phosphorylation; Receptor; Repeat; Signal; Transmembrane; Wnt signaling pathway.

FT SIGNAL 1 30 Potential.

FT CHAIN 31 1614 Low-density lipoprotein receptor-related protein 5.

FT DOMAIN 31 1383 Extracellular (Potential).

FT TRANSMEM 1384 1406 Potential.

FT DOMAIN 1407 1614 Cytoplasmic (Potential).

FT DOMAIN 74 117 LDL-receptor class B 1.

FT DOMAIN 119 160 LDL-receptor class B 2.

FT DOMAIN 162 204 LDL-receptor class B 3.

FT DOMAIN 206 245 LDL-receptor class B 4.

FT DOMAIN 247 288 LDL-receptor class B 5.

FT DOMAIN 294 336 EGF-like 1.

FT DOMAIN 384 425 LDL-receptor class B 6.

FT DOMAIN 427 468 LDL-receptor class B 7.

FT DOMAIN 470 512 LDL-receptor class B 8.

FT DOMAIN 514 557 LDL-receptor class B 9.

FT DOMAIN 600 640 EGF-like 2.

FT DOMAIN 686 727 LDL-receptor class B 10.

FT DOMAIN 729 770 LDL-receptor class B 11.

FT DOMAIN 772 813 LDL-receptor class B 12.

FT DOMAIN 815 852 LDL-receptor class B 13.

FT DOMAIN 855 896 LDL-receptor class B 14.

FT DOMAIN 901 941 EGF-like 3.

FT DOMAIN 988 1033 LDL-receptor class B 15.

FT DOMAIN 1035 1076 LDL-receptor class B 16.

FT DOMAIN 1078 1121 LDL-receptor class B 17.

FT DOMAIN 1123 1164 LDL-receptor class B 18.

FT DOMAIN 1165 1206 LDL-receptor class B 19.

FT DOMAIN 1212 1253 EGF-like 4.

FT DOMAIN 1257 1295 LDL-receptor class A 1.

FT DOMAIN 1296 1332 LDL-receptor class A 2.

FT DOMAIN 1334 1370 LDL-receptor class A 3.

FT SITE 1499 1503 PPPSP motif.

FT SITE 1605 1609 By similarity.

FT DISULFID 298 309 By similarity.

FT DISULFID 305 320 By similarity.

FT DISULFID 322 335 By similarity.

FT DISULFID 604 615 By similarity.

FT DISULFID 611 624 By similarity.

FT DISULFID 626 639 By similarity.

FT DISULFID 905 916 By similarity.

FT DISULFID 912 925 By similarity.

FT DISULFID 927 940 By similarity.

FT DISULFID 1216 1227 By similarity.

FT DISULFID 1223 1237 By similarity.

FT DISULFID 1239 1252 By similarity.

FT DISULFID 1258 1272 By similarity.

FT DISULFID 1265 1285 By similarity.

FT DISULFID 1279 1294 By similarity.

FT DISULFID 1297 1309 By similarity.

FT DISULFID 1304 1322 By similarity.

FT DISULFID 1316 1331 By similarity.

FT DISULFID 1335 1347 By similarity.

FT DISULFID 1342 1360 By similarity.

FT DISULFID 1354 1369 By similarity.

FT CARBOHYD 92 92 N-linked (GlcNAc...) (Potential).

FT CARBOHYD 137 137 N-linked (GlcNAc...) (Potential).

FT CARBOHYD 445 445 N-linked (GlcNAc...) (Potential).

```

FT CARBOHYD 498 498 N-linked (GlcNAc...) (Potential)
FT CARBOHYD 704 704 N-linked (GlcNAc...) (Potential)
FT CARBOHYD 877 877 N-linked (GlcNAc...) (Potential)
FT CONFLICT 220 220 R -> H (in Ref. 2)
FT CONFLICT 1520 1520 I -> S (in Ref. 3)
FT CONFLICT 1553 1553 I -> T (in Ref. 3)
SQ SEQUENCE 1614 AA; 178896 MW; 911EB288471752C5 CRC64;

Query Match 94.5%; Score 8259.5; DB 1; Length 1614;
Best Local Similarity 93.9%; Pred. No. 0;
Matches 1519; Conservative 41; Mismatches 51; Indels 7; Gaps 2;

QY 1 MEAAP---PGPEWPLLLLLLLLLLALCGCPAPAAASPLLFANRRDRLVLDAGGVKLESTI 57
DB 1 METAPTRAPPPPPPLLLLVLYCSL-----VPAASPLLFANRRDRLVLDAGGVKLESTI 56
QY 58 VVSGLEDAADVDFQSKGAVYMTDVSEBAIKQTYLNOTGAQVONVVISGLVSPDGLACDW 117
DB 57 VASGLEDAADVDFQSKGAVYMTDVSEBAIKQTYLNOTGAQAQNVIVISGLVSPDGLACDW 116
QY 118 VGKLYWTDSETNRIEVANLNGTSRKVLFWODLQDPAIRALDPAHYMYWTDWGETPRIE 177
DB 117 VGKLYWTDSETNRIEVANLNGTSRKVLFWODLQDPAIRALDPAHYMYWTDWGEAPRIE 176
QY 178 RAGMDGSTRKIIIVSDIYWPNGLATIDLEQKLYWADAKLSFIHRANLDGSPQKRVGSL 237
DB 177 RAGMDGSTRKIIIVSDIYWPNGLATIDLEQKLYWADAKLSFIHRANLDGSPQKRVGSL 236
QY 238 THPFALTLSGDTLWTDWQTRSIIHACNKRGTGGKKEILSALYSPMDIQVLQSOERPPFHT 297
DB 237 THPFALTLSGDTLWTDWQTRSIIHACNKRGTGGKKEILSALYSPMDIQVLQSOERPPFHT 296
QY 298 RCEBDNGCSHLCLLSSEPEYTCACPTGVOLQNGRTCKAGAEVLLIARTRILSL 357
DB 297 PCEBDNGCSHLCLLSSEPEYTCACPTGVOLQNGRTCKAGAEVLLIARTRILSL 356
QY 358 DTPDFTDVLQVDDIRHAIADYDPLEGVYVWTDDEVAIRRAYLDGSGAOTLVNTEIND 417
DB 357 DTPDFTDVLQVDDIRHAIADYDPLEGVYVWTDDEVAIRRAYLDGSGAOTLVNTEIND 416
QY 418 PDGIAVDWAAVNLWTDGTDRIEVTRNLGTSRKILVSEDLDEPRAIALHPVGLMYWTD 477
DB 417 PDGIAVDWAAVNLWTDGTDRIEVTRNLGTSRKILVSEDLDEPRAIALHPVGLMYWTD 476
QY 478 WGENPKTECANLQGERVLVNASIGWPNGLALDLOEGKLYWDAKTDKIEVINVDGTR 537
DB 477 WGENPKTECANLQGERVRLVNTSLGPNGLALDLOEGKLYWDAKTDKIEVINVDGTR 536
QY 538 RTLEDKLPHPFGFTLLGDFIYTDWQRRSIRVHKVKASRDVVIDQLPDLMLKAVNVA 597
DB 537 KTLLEDKLPHPFGFTLLGDFIYTDWQRRSIRVHKVKASRDVVIDQLPDLMLKAVNVA 596
QY 598 KVGVTNFCADNRGGCSHLCPFTPHATRCGCPIGLELLSDMKTCIVPRAFLVFTSRAIHR 657
DB 597 KVGVTNFCADNRGGCSHLCPFTPHATRCGCPIGLELLSDMKTCIPEAFVFTSRATTHR 656
QY 658 ISLETNNDVAILPTGVKEASALDPDVSNNHIYTDVLSLXTISRAFWNGSSVEHVVERGL 717
DB 657 ISLETNNDVAILPTGVKEASALDPDVSNNHIYTDVLSLXTISRAFWNGSSVEHVVERGL 716
QY 718 DYPEGMAVDMWGNKLYWADTGTNRIEVARLDGQFQVLRDLNPRSLALDPTKGYIYW 777
DB 717 DYPEGMAVDMWGNKLYWADTGTNRIEVARLDGQFQVLRDLNPRSLALDPTKGYIYW 776
QY 778 TEWGGKPRIVAFWGTGNTMTLVKVGANDLTIDYADQRLYWTDLTNMISSNMLGQE 837
DB 777 TEWGGKPRIVAFWGTGNTMTLVKVGANDLTIDYADQRLYWTDLTNMISSNMLGQE 836
QY 838 RVVIADLLPFPFGITQYSDIYTWDMNLHSIERADKTSGRNRTLIQGLHDFWMDILVFS 897
DB 837 RVVIADLLPFPFGITQYSDIYTWDMNLHSIERADKTSGRNRTLIQGLHDFWMDILVFS 896
QY 898 SRQGLNDCHMNNQCQGLCLAIPEGHRCGCASHYTLDPSSRNCSPPTFTLLFSQKSAIS 957

```

```

DB 897 SRQGLNDCHMNNQCQGLCLAIPEGHRCGCASHYTLDPSSRNCSPPTFTLLFSQKSAIS 956
QY 958 RMIPDDQHSPLDILPLHGLRNKKAIDYDPLDKFIYWDGRQNKRAKADGTQPPVLTSL 1017
DB 957 RMIPDDQHSPLDILPLHGLRNKKAIDYDPLDKFIYWDGRQNKRAKADGTQPPVLTSL 1016
QY 1018 QGQNPDRQPHDLSIDISRTLFTWCEATNTINVRHLSGEAMGVLRGDRDKPRRAIVNAE 1077
DB 1017 QLSFDRQPHDLSIDISRTLFTWCEATNTINVRHLSGEAMGVLRGDRDKPRRAIVNAE 1076
QY 1078 RGYLYFTNMQDRAAKIERAALDGTREVLFTGLIRPVVALVVDNLTGLKFWVDADLKR 1137
DB 1077 RGYLYFTNMQDRAAKIERAALDGTREVLFTGLIRPVVALVVDNLTGLKFWVDADLKR 1136
QY 1138 SCDLSGANRLTLEDANIVQPLGLTILGKHLWIDRQQQMIERVEKTTGDKRTIQGRVAH 1197
DB 1137 SCDLSGANRLTLEDANIVQPLGLTILGKHLWIDRQQQMIERVEKTTGDKRTIQGRVAH 1196
QY 1198 LTGTHAVEVSELEFSAPHCARDNGGCSHICIAKGDGTTPRCSCPVLHVLQNLITCGEPP 1257
DB 1197 LTGTHAVEVSELEFSAPHCARDNGGCSHICIAKGDGTTPRCSCPVLHVLQNLITCGEPP 1256
QY 1258 TCSPDQACATGEIDCIPGAWRCDFPECDQSDDEGCPVCSAAQFPCCARGOCVDLRRC 1317
DB 1257 TCSPDQACATGEIDCIPGAWRCDFPECDQSDDEGCPVCSAAQFPCCARGOCVDLRRC 1316
QY 1318 DGEADQDRSDEADCAICLPNQPRCASGQCVLIKQCCDSDPDICIDGSDLMCEITKPPS 1377
DB 1317 DGEADQDRSDEADCAICLPNQPRCASGQCVLIKQCCDSDPDICIDGSDLMCEITKPPS 1376
QY 1378 DDPASASATGPVIGIILSLFVMGVYVFCORVVCORVYAGANGPPHVEYVGTGHPVPLNF 1437
DB 1377 DDPASASATGPVIGIILSLFVMGVYVFCORVVCORVYAGANGPPHVEYVGTGHPVPLNF 1436
QY 1438 IAPGSGHQGPPTGIACGKSMSSVSLMGRCGVPLDYDRNHVTGASSSSSSSTKATLYPPI 1497
DB 1437 IAPGSGHQGPPTGIACGKSMSSVSLMGRCGVPLDYDRNHVTGASSSSSSSTKATLYPPI 1496
QY 1498 LNPPSPATDPSLVNMDYSSNIPATAPRVPIIIRGMAPPTTPTCSTDVCDSDYSASRW 1557
DB 1497 LNPPSPATDPSLVNMDYSSNIPATAPRVPIIIRGMAPPTTPTCSTDVCDSDYSASRW 1556
QY 1558 KASKYVLDLNSDSDPYPPPTPHSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615
DB 1557 KASKYVLDLNSDSDPYPPPTPHSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1614

RESULT 4
Q8AYF1 PRELIMINARY; PRT; 1605 AA.
AC Q8AYF1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Lipoprotein receptor-related protein 5.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22194715; PubMed=12204281; DOI=10.1016/S0925-4773(02)00205-8;
RA Houston D.W., Wylie C.;
RT "Cloning and expression of Xenopus Lrp5 and Lrp6 genes.";
RL Mech. Dev. 117:337-342(2002).
DR EMBL; AF508960; AAN0906.1; -.
DR HSSP; P01130; 1D2J.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR006210; IEGF.

```


"Isolation and Characterization of LRP6, a novel member of the low density lipoprotein receptor gene family.";
 Biochem. Biophys. Res. Commun. 248:879-888(1998).
 [2]
 FUNCTION, AND DOMAINS REQUIRED FOR INTERACTION WITH DKK1 AND DKK2.
 PubMed=11357136; DOI=10.1038/35077108;
 RA Mac B., Wu W., Li Y., Hoppe D., Stannek P., Glinka A., Niehrs C.;
 RT "LDL-receptor-related protein 6 is a receptor for Dickkopf proteins.";
 RL Nature 411:321-325(2001).
 RN [3]
 RN INTERACTIONS WITH FZD5; DKK1 AND DKK2.
 RP PubMed=12857724; DOI=10.1074/jbc.M300191200;
 RA Caricasole A., Ferraro T., Iacovelli L., Barletta E., Caruso A.,
 RA Melchiorri D., Terstappen G.C., Nicoletti F.;
 RT "Functional characterization of WNT7A signaling in PC12 cells:
 RT interaction with A FZD5 x LRP6 receptor complex and modulation by
 RT Dickkopf proteins.";
 RL J. Biol. Chem. 278:37024-37031(2003).
 RN [4]
 RN REVIEW
 RP PubMed=15084453; DOI=10.1242/dev.01117;
 RX He X., Semenov M., Tamai K., Zeng X.;
 RA "LDL receptor-related proteins 5 and 6 in Wnt/[beta]-catenin
 RT signaling: Arrows point the way.";
 RL Development 131:1663-1677(2004).
 CC -1- FUNCTION: Essential for the Wnt/beta catenin signaling pathway,
 CC probably by acting as a coreceptor together with Frizzled for Wnt.
 CC Specific high-affinity receptor for DKK1 and DKK2, but not DKK3.
 CC The interaction with DKK1 blocks LRP6-mediated Wnt/beta catenin
 CC signaling via LRP6 removal via Kremen proteins-mediated
 CC endocytosis.
 CC -1- SUBUNIT: Interacts with FZD5. Essential component of the Wnt7
 CC receptor complex. Wnt7A interacts with the LRP6/FZD5 complex. This
 CC interaction is antagonized by DKK1 and DKK3.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
 CC -1- TISSUE SPECIFICITY: Widely co-expressed with LRP5 during
 CC embryogenesis and in adult tissues.
 CC -1- DOMAIN: The WNT-EGF-like domains 1 and 2 are required for the
 CC interaction with Wnt-frizzled complex. The WNT-EGF-like domains 3
 CC and 4 are required for the interaction with DKK1.
 CC -1- SIMILARITY: Belongs to the LDLR family.
 CC -1- SIMILARITY: Contains 4 EGF-like domains.
 CC -1- SIMILARITY: Contains 3 LDL-receptor class A domains.
 CC -1- SIMILARITY: Contains 20 LDL-receptor class B domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF074264; AAC33006.1; -;
 CC PIR; JE0272; JE0272.
 CC HSP; P01130; 1AJJ.
 CC Genew; HGNC:6698; LRP6.
 CC MIM; 603507; -;
 CC GO; GO:0016021; C: integral to membrane; TAS.
 CC GO; GO:0005886; C: plasma membrane; TAS.
 CC GO; GO:0005515; F: protein binding; TAS.
 CC InterPro; IPR000742; EGF_2.
 CC InterPro; IPR006209; EGF_like.
 CC InterPro; IPR006210; IEGF.
 CC InterPro; IPR002172; LDL_receptor_A.
 CC InterPro; IPR000033; Ldl_receptor_rep.
 CC Pfam; PF00008; EGF_4.
 CC Pfam; PF00057; Ldl_recept_a; 3.
 CC Pfam; PF00058; Ldl_recept_b; 20.
 CC PRINTS; PR00261; LDLRECEPTOR.
 CC SMART; SM00181; EGF_4.
 CC SMART; SM00192; LDLa; 3.
 CC SMART; SM00135; LY; 20.

DR	PROSITE; PS00022; EGF_1; FALSE_NEG.			
DR	PROSITE; PS01186; EGF_2; 1.			
DR	PROSITE; PS00026; EGF_3; FALSE_NEG.			
DR	PROSITE; PS01209; LDLRA 1; 3.			
KW	Developmental protein; LDLR; EGF-like domain; Endocytosis; Glycoprotein; Receptor; Repeat; Signal; Transmembrane; Wnt signaling pathway.			
FT	SIGNAL 1 19	Potential.		
FT	CHAIN 20 1613	Low-density lipoprotein receptor-related protein 6.		
FT	DOMAIN 20 1370	Extracellular (Potential).		
FT	TRANSMEM 1371 1393	Potential.		
FT	DOMAIN 1394 1613	Cytoplasmic (Potential).		
FT	DOMAIN 63 105	LDL-receptor class B 1.		
FT	DOMAIN 107 148	LDL-receptor class B 2.		
FT	DOMAIN 150 192	LDL-receptor class B 3.		
FT	DOMAIN 194 235	LDL-receptor class B 4.		
FT	DOMAIN 237 276	LDL-receptor class B 5.		
FT	DOMAIN 282 324	EGF-like 1.		
FT	DOMAIN 372 413	LDL-receptor class B 6.		
FT	DOMAIN 415 456	LDL-receptor class B 7.		
FT	DOMAIN 458 500	LDL-receptor class B 8.		
FT	DOMAIN 502 540	LDL-receptor class B 9.		
FT	DOMAIN 543 586	LDL-receptor class B 10.		
FT	DOMAIN 588 628	EGF-like 2.		
FT	DOMAIN 674 715	LDL-receptor class B 11.		
FT	DOMAIN 717 758	LDL-receptor class B 12.		
FT	DOMAIN 760 801	LDL-receptor class B 13.		
FT	DOMAIN 803 840	LDL-receptor class B 14.		
FT	DOMAIN 843 884	LDL-receptor class B 15.		
FT	DOMAIN 889 930	EGF-like 3.		
FT	DOMAIN 977 1024	LDL-receptor class B 16.		
FT	DOMAIN 1026 1067	LDL-receptor class B 17.		
FT	DOMAIN 1069 1112	LDL-receptor class B 18.		
FT	DOMAIN 1114 1155	LDL-receptor class B 19.		
FT	DOMAIN 1157 1197	LDL-receptor class B 20.		
FT	DOMAIN 1203 1244	EGF-like 4.		
FT	DOMAIN 1248 1286	LDL-receptor class A 1.		
FT	DOMAIN 1287 1323	LDL-receptor class A 2.		
FT	DOMAIN 1325 1361	LDL-receptor class A 3.		
FT	SITE 1487 1491	PPSP motif.		
FT	SITE 1604 1608	By similarity.		
FT	DISULFID 286 297	By similarity.		
FT	DISULFID 293 308	By similarity.		
FT	DISULFID 310 323	By similarity.		
FT	DISULFID 592 603	By similarity.		
FT	DISULFID 599 612	By similarity.		
FT	DISULFID 614 627	By similarity.		
FT	DISULFID 893 904	By similarity.		
FT	DISULFID 900 914	By similarity.		
FT	DISULFID 916 929	By similarity.		
FT	DISULFID 1207 1218	By similarity.		
FT	DISULFID 1214 1228	By similarity.		
FT	DISULFID 1230 1243	By similarity.		
FT	DISULFID 1249 1263	By similarity.		
FT	DISULFID 1256 1276	By similarity.		
FT	DISULFID 1270 1285	By similarity.		
FT	DISULFID 1288 1300	By similarity.		
FT	DISULFID 1295 1313	By similarity.		
FT	DISULFID 1307 1322	By similarity.		
FT	DISULFID 1326 1338	By similarity.		
FT	DISULFID 1333 1351	By similarity.		
FT	DISULFID 1345 1360	By similarity.		
FT	CARBOHYD 42 42	N-linked (GlcNAc. . .)	(Potential).	
FT	CARBOHYD 81 81	N-linked (GlcNAc. . .)	(Potential).	
FT	CARBOHYD 281 281	N-linked (GlcNAc. . .)	(Potential).	
FT	CARBOHYD 433 433	N-linked (GlcNAc. . .)	(Potential).	
FT	CARBOHYD 486 486	N-linked (GlcNAc. . .)	(Potential).	
FT	CARBOHYD 592 592	N-linked (GlcNAc. . .)	(Potential).	
FT	CARBOHYD 859 859	N-linked (GlcNAc. . .)	(Potential).	
FT	CARBOHYD 865 865	N-linked (GlcNAc. . .)	(Potential).	
FT	CARBOHYD 926 926	N-linked (GlcNAc. . .)	(Potential).	
FT	CARBOHYD 1039 1039	N-linked (GlcNAc. . .)	(Potential).	

SQ	SEQUENCE	1613 AA; 180442 MW; 4BC1141E395D8B5B CRC64;
Query Match	70.4%; Score 6153.5; DB 1; Length 1613;	
Best Local Similarity	70.2%; Pred. No. 0;	
Matches 1134; Conservative	205; Mismatches 256; Indels 21; Gaps 12;	
Qy	15 LLLLLLALCGCPAPAAASPLLLPANRRDRLVLDAGGVKLESTIVVSGLEDAADVPFQSK 74	
Db	4 VLRSLLACSFVLLRAA-PLLLYANRRDLRLVDATNGKENATIVVGGLEDAADVPFVFSH 62	
Qy	75 GAVYTWDSBEAIIKOTYLNOGAOVNVVIGSLVSPDGLACDWGKKLYWTDSTNRIEV 134	
Db	63 GLIYWSVSEBAIKRTFENKT-ESQNVVVSGLLSPDGLACDWLGEKKLYWTDSETNRIEV 121	
Qy	135 ANLNGTSRKVLFWQDLQOPRAIALDPDAHGYMYWTDWGETPRIERAGMDGSTRKIIVDSDI 194	
Db	122 SNLGSRLKVLFWQDLQOPRAIALDPSSGFYWTDWGEVPKIERAGMDGSSRFIIINSEI 181	
Qy	195 YWPNGLIIDLEQKLYWADAKLSIFHRANLDGSPROKVVGEGLTHPFAALTUSGDTLIWTD 254	
Db	182 YWPNGLIIDLEQKLYWADAKLNFHKSNDLGTNRQAVVKGSLPHFPFALTIFEDILIYWD 241	
Qy	255 WQTESIHACNKRCTGKKEIILSALYSMDIQVLQOEROPFHTCEBDNGCSCHLCLSP 314	
Db	242 WSTHSIIACNKYTGELUREIHSIDIFSPMDIHAFSQOQPNAATPCGIDNGCSCHLCLMSP 301	
Qy	315 SEPFYTCACPTGVOLQNGRTCKAGAEVLLARLTDLRISLDTPOFTDIVLQVDDIRH 374	
Db	302 VKPFYQACPTGVKLENGKTCKGATELLLLARLTDLRISLDTPOFTDIVLQLEDIRH 361	
Qy	375 AIAIDYDPLEGYVYWTDEVAIRAYLDGSGAOTLVNTEINDPDGIADVWVARNLYWTD 434	
Db	362 AIAIDYDPEGYIYWTDEVAIRARRSPIDGSGSFVVTQAIAHPDGIADVWVARNLYWTD 421	
Qy	435 TGTDRIEVTRNGTSRKILVSEDLDEPRAIALHPVWGLMYWTDWGENPKIECANLDQER 494	
Db	422 TGTDRIEVTRNGTMRKILVSEDLDEPRAIYLDPMVGYMYWTDWGEIPIKIERAALDGSDR 481	
Qy	495 RVLVNASLGNWGLALDLOEKLWYGAOKTDKIEVINVDGTPKRTLLLEDKLPHPFGTLL 554	
Db	482 VLVNTSLGNWGLALDYDEKIIYGAOKTDKIEVMWTDGTRRVLVEDKLPHPFGTLL 541	
Qy	555 GDFIYTWDRQRRSIRVHKVKSARDVIIDQLPDLMLKANVAVVGTNPNPCADRNGGCSH 614	
Db	542 GDYVYTWDRQRRSIRVHKVKSAREVIIDQLPDLMLKATNVHVRVIGSNPCABENGCGSH 601	
Qy	615 LCFPTPHATRCGCGIGLELLSDMKTCIVPEAFIVFTSRAAHRISLETNNNDVAIPLTGV 674	
Db	602 LCLYRPOGLRCACPIGFELISDMKTCIVPEAFLLFSRRADIRISLETNNNNVAIPLTGV 661	
Qy	675 KEASALDPDVSNHHIYWTDVSLKTIISRAFPMNGSSVEHVVEFGLDYPEGMAVDWVGKLYW 734	
Db	662 KEASALDPDVTDNRIYWTDISLKTISRAFPMNGSALEHVVEFGLDYPEGMAVDWVGKLYW 721	
Qy	735 ADTGTRNIEVARLDQGRQVILVWRDLNPRSLADPTKGYIYTWEGGKPRIVAFMDGT 794	
Db	722 ADTGTRNIEVSKLDQGRQVILVWKDLSPRALADPAEGFMYWTEWGGKPKIDRAAMDGS 781	
Qy	795 NCMTLVKVGANDLTIIDYADORLYWTDLDTNMWLESSNMLQOERVVADDLPHFPGLTQY 854	
Db	782 ERTTLVPNVGVRANGLTIIIDYAKRRLYWTDLDTNLISSNMLGLNREVADDLPHFPGLTQY 841	
Qy	855 SDYIYWTDNWLNHSTRADKTSGRNRTLIQGLHDFVMDILVPHSSROGLNDCMNNNGCG 914	
Db	842 QDYIYWTDNWSSRSTRANKTSGRNRTLIQGLHDFVMDILVPHSSROGWNECASNGHCS 901	
Qy	915 QLCILAIP-GGHRCCGASHYTLDPSSRNCSPPTTFLFSQKSAISRMIPDDQHSPLIPL 973	
Db	902 HLCIAVPVGGVCGPAHYSLNADNRCTCSAFTTFLFSQKSAINRWVIDEQSQSDIILPI 961	
Qy	974 HGLRNVAIDYDPLDKFIYWDGQRN-IKRAKDGTQPF-VLTSLSQGNPDPRPHDLIS 1031	
Db	962 HSLRNVAIDYDPLDKOLYWDSDRQNMIRKAQEDGSGQFTVVWSSVPSQNLIEIQPYDLIS 1021	

Qy	1032 DIYSRTLFTWCEATNTINVHRLSGEAMGVVLRGDRDKPRATVVAERGLYLTNNQDRAA 1091	
Db	1022 DIYSRYIYWCETATNVINVRFLDGRSVGVVLKGGQDRPRALVNVNPKGYWYFTNLQERSP 1081	
Qy	1092 KIERAALDGTERTEREVLFTGLIRPVVALVVDNTLGLKFWVDADLKRIESCGLSGANRLTLED 1151	
Db	1082 KIERAALDGTERTEREVLFFSGLSKPTIALADSLRGLKFWADSLRRIESDLSGANRIVLED 1141	
Qy	1152 ANIYQPLGLTILGLKHLWIDRQQQMIERVEKTDGKTRIOGRVAHLTGIIHAREVSELEE 1211	
Db	1142 SNILQPVGLIVFENWLWIDKQQQMIKIDMTGREGRTKVOARLAQSLDIHAYVELNLQE 1201	
Qy	1212 FSAHPCARDNGGCSHICIAKGDGTTPRCSCPVLHLLQNLQLTCGPPPTCPQFPACATGEI 1271	
Db	1202 YRQHPCADNGGCSHICLVKGDGTTTRCSCPMHLVLLQDELSCGEPPTCSPOQFTCTGEI 1261	
Qy	1272 DCIPGAWRCDFPECDQSDSEEGCPVCSAAQPPCARGOCVDLRLRCDEACQDRSDEAD 1331	
Db	1262 DCIPVAWRCDFTECDHSDDELNCPCVSESQFCASGQCIDGALRCNGDANCQDKSDEKN 1321	
Qy	1332 CDAICLNPFRCSAGCQVLKQOQCDSPDCIDGSDLMCEITKPPSDDSPAHSSAIGPVI 1391	
Db	1322 CEVLCLLDQFRCAAGQCIGKHKCDHNVDCSKDELDCYTEEP---AQAATNVGSVI 1378	
Qy	1392 GIILSLFVMGGVYFVQCQRYAGANGPPPHHEY-VSGTTPHVPLNFTIAPGSGHQGPFTG 1450	
Db	1379 GVIVTIFVSGTVYFICQMLCPKMGDGTMTNDVYVHGPPASVPLGYVPHPSLGSPLG 1438	
Qy	1451 IACKSMSSVSLMGRGGVPLYDRNHVTGASSSSSSSTKATLYPPIINPPPPSPATDPSL 1510	
Db	1439 MSRGKMISSLSIMGSSGPP-YDRAHVTGASSSSSSSTKGYFPAILNPPPPSPATERSH 1497	
Qy	1511 YNMDFYSSNIPATAR--PYRPYIIRGMAPPTTCTDVCDSYDSASR-----WKASKY 1563	
Db	1498 YTFEGYSSNSPSYRHSYRPSYRHPFAPPTTCTDVCDSYDSASRMTSVATAKYT 1557	
Qy	1564 LDLNSDDPPPPPTPHSQVLSAE---DSCPPSPATERSY-FHLFPPPPSPCTDSS 1615	
Db	1558 SDLNYDSEPPPPPTPRSQVLSAENYSCPPSPYTERSYSHHLYPPPPSPCTDSS 1613	

RESULT 6

LRP6 MOUSE

ID_LRP6_MOUSE STANDARD; PRT; 1613 AA.

AC O89572;

DT 05-JUL-2004 (Rel. 44, Created)

DT 05-JUL-2004 (Rel. 44, Last sequence update)

DT 05-JUL-2004 (Rel. 44, Last annotation update)

DE Low-density lipoprotein receptor-related protein 6 precursor.

GN Name=Lrp6;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

ON NCBI_TaxID=10090;

OX [1]

PC SEQUENCE FROM N.A.

RP STRAIN=BALB/c; TISSUE=Liver;

RC MEDLINE=98369644; PubMed=9704021; DOI=10.1006/bbrc.1998.9061;	
RA Brown S.D., Twells R.C., Hey P.J., Cox R.D., Levy E.R., Soderman A.R.,	
RA Metzker M.L., Caskey C.T., Todd J.A., Hess J.F.,	
RT "Isolation and characterization of LRP6, a novel member of the low	
RT density lipoprotein receptor gene family."	
RL Biochem. Biophys. Res. Commun. 248:879-888(1998).	
RN [2]	
RP SEQUENCE FROM N.A.	
RC STRAIN=C57BL/6; TISSUE=Brain;	
RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;	
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,	
RA Klausner R.G., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,	
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,	
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,	
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,	

Db 4 VLRSLACFCVLLRAA-PLLYANRRDLRLVDAATNGKENATIVVGGLEDAADAAVDFVGH 62
Qy 75 GAVYWDVSEBAIKOTLYNOTGAAYQNVVIGLVSPDGLACDWGKLYWTDSETRIEV 134
Db 63 GLIYWDVSEBAIKETENKS-ESVQNVVSGLLSPDGLACDWLGEKLYWTDSETRIEV 121
Qy 135 ANLNGTSKRVLPWQDQFRAIALDPAHGYWYTDWGETPRIERAGMDGSTRKIIVDSI 194
Db 122 SNLDGSLKRVLPWQDQFRAIALDPSGGMFWYTDWGEVPKIERAGMDGSSRFVIINTEI 181
Qy 195 YWPNGLTIDLEOKLYWADAKLSPFTHRANLDGSPROKVVESGLTHPPALTISGDTLYWTD 254
Db 182 YWPNGLTIDYQERKLYWADAKLNFTHKSNLDGTRQAVVKGSLPHFPALTIFEDTLYWTD 241
Qy 255 WQTSIHACNKRKTGGKKEILSALYSPMDIOVLQSERQPFPHTRCEEONGGCSHLCLLSP 314
Db 242 WNTSILACNKTGEGLEIHSNTPSPMDIHAFSQORPNATNPGCINDGGCSHLCLLSP 301
Qy 315 SEPFTYCAPTGVQDNGRTCKAGAERVLRLARRTDLRRISLDTPTFDIVLQVDDIRH 374
Db 302 VKPFYQACPTGVKLMENGKTKCDGATELLARRTDLRRISLDTPTFDIVLQLEDIRH 361
Qy 375 AIAIDYDPLEGVVYTDDEVAIRRAYLDGSGAQLVNTENDPGIADVWVARNLYWTD 434
Db 362 AIAIDYDVEGYIYWTDDEVAIRRSFIDGSGSQFVVTQAIAHPDGIADVWVARNLYWTD 421
Qy 435 TGTDRIEVTRNGTSRKILVSEDLDEPRAIALHPVMGLMYTMDGENPKIECANLDGOER 494
Db 422 TGTDRIEVTRNGTWKILISEDLDEEPRAIVLDPWGYWYTDWGEIPKIERAALDGSDR 481
Qy 495 RVLVNASLWPNGLALDLOEGKLYWGDARTKIEVINVDGTRKRTLLBKLPHPFGFTLL 554
Db 482 VVLVNTSLGWPNGLALDYDEGTIYWGDAKTKEIEMVNTDGTGRRVLVEDKIPHPFGFTLL 541
Qy 555 GDFIYWDWQRRSERVHKVVASRDVIIDQLPDLMLKANVAVKVVGNVPCADNRGGCSH 614
Db 542 GDIYVWDWQRRSERVHKVSAEREVIIDQLPDLMLKATSVHRVIGSNPCAEONGGCSH 601
Qy 615 LCFTEPHATRCGPGLLELLSDMTKCIVPEAPLFTVTSAAIHRISLENNNDVAIPLTVG 674
Db 602 LCLYRPOGLRCACPGLIGELIGDMTKCIVPEAPLFTVTSAAIHRISLENNNDVAIPLTVG 661
Qy 675 KEASALDPDVNNHIYWDVSLKTSIRAFMNGSSVHVVEFGLDYPEGMADVWNGKLYW 734
Db 662 KEASALDPDVNNHIYWDVSLKTSIRAFMNGSALEHVVEFGLDYPEGMADVWNGKLYW 721
Qy 735 ADTGTRIEVARLDGQFQVILVWRDLNPRSLALDPTKGYIYTWEGGKPRIVAFMDGT 794
Db 722 ADTGTRIEVARLDGQFQVILVWRDLNPRSLALDPAEGFMWTEWGGKPIDRAAMDGS 781
Qy 795 NCMTLVKVGANDLTIDYADORLYWTDLDTNMISSNMLQERWVIADDLPHFPGLTOY 854
Db 782 ERTTLVNVVGRANGLTIDYAKRUYWTDLDTNLISSDMLGAINREVATDDLPHFPGLTOY 841
Qy 855 SDIYIYTDWNLHSIERADKTSGRNRTIQCHLDFVMDILVFHSSRQDGLNDCMHNQCG 914
Db 842 QDIYIYTDWNSRISERANKTSQNRITIQCHLDFVMDILVFHSSRQAGWNECASNGHCS 901
Qy 915 QLCIAIP-GGHCRCASHYLDPSRNCSPPTFLFSQKSAISRMIPDDQHSPLILPL 973
Db 902 HLCIAVPVGGFCVCPAHYSLNADNRTCSAPSTFLFSQKSAINRWVIDEQSPDIILPI 961
Qy 974 HGLRNVKAIYDPLDKIYWVWGRQN-IKRAKDGTQPP-VLTSLSQGNPDPRPHLISI 1031
Db 962 HSLRNVKAIYDPLDKIYWVWGRQNSIRKAHEPDGGGFNVVANSVANQNLIEQFYDLIS 1021
Qy 1032 DIYSRTLPWTEATNTNVHRLSGEAMGVLRGRDKPRVIAVNAERGLYFTWMDRAA 1091
Db 1022 DIYSRYIYTWTEATNTVDVTRLDGRSVGVVLKGGQDRPRVIAVNAERGLYFTWMDRAA 1081
Qy 1092 KIERAALDGTREVLFTGLIRPVVALVVDNTLGLKFWVDADLKRIESCDSGANRLTIED 1151
Db 1082 KIERAALDGTREVLFTGLIRPVVALVVDNTLGLKFWVDADLKRIESCDSGANRLTIED 1141

Qy 1152 ANIVQPLGLTILGKHLWYIDRQOQMIERVEKTTGTRTRIOGRVAHLGTIHAVERVSLEE 1211
Db 1142 SNILQPLGLTIFENWLYWIDRQOQMIERVEKTTGTRTRIOGRVAHLGTIHAVERVSLEE 1201
Qy 1212 FSAHPCARDNGCGSHICIAKGDGTPRCPCVHLVLLQNLATCGEPPTCSPDQFACATGEI 1271
Db 1202 YRQHPCAQDNGCGSHICIAKGDGTPRCPCVHLVLLQNLATCGEPPTCSPDQFACATGEI 1261
Qy 1272 DCIFGAWRCDFPFCDDOSDEEGPCVCSAAQPPCARGOCVLDRLRLCDGEADQCDSDEAD 1331
Db 1262 DCIFPVMWRCDFPFCDDOSDEEGPCVCSAAQPPCARGOCVLDRLRLCDGEADQCDSDEAD 1321
Qy 1332 CDALCLNPQFRCAGQCQVLIKQOQDPSFDCIDGSDLMCEITKPPSDOSPAHSSAIGPVI 1391
Db 1322 CEVLCLDLOFCANGQCVKHKCDHSDVDCSDRDELDCYTEEP---APQNTVTVGSVI 1378
Qy 1392 GIILSLFVMGVGVYVFCQVRCVQYAGANGPPHEHYVSGTP-HVPLNFIAPGSGORHPTG 1450
Db 1379 GVIITFVSGTIFTCQMLCPRMKGDGTWNTDYYVHSPASVPLGYVPHPSLSGSLPG 1438
Qy 1451 IACGKSMSSVSLMGRGVPLYDRNHVTGASSSSSSSTKATLYPPIILNPPSPATDPSL 1510
Db 1439 MSRGKSMSSVSLMGRGVPLYDRNHVTGASSSSSSSTKATLYPPIILNPPSPATDPSL 1497
Qy 1511 YNMDWFSSNIPATAR--PYRPIIIRGMWAPPTPCSTDVCDSDYSASR----WKASKYV 1563
Db 1498 YTMFEGYSSNSPSTHRSYRPSYRHPFAPPTTTCSTDVCDSDYAPSRMRTSVATAKGYT 1557
Qy 1564 LDLNSDSDYPPPPPHPSQYLSAE---DSCPPSPATERSY-FHLPPPPSPCTDSS 1615
Db 1558 SDVNYDEPVPPTTPRSQYLSAEVSCPPSPYTERSYSHHLYPPPPSPCTDSS 1613

RESULT 7

Q8AYFO PRELIMINARY; PRT; 1613 AA.
AC Q8AYFO; SEQUENCE FROM N.A.
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Lipoprotein receptor-related protein 6.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
NC NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22194715; PubMed=12204281; DOI=10.1016/S0925-4773(02)00205-8;
RA Houston D.W., Wylie C.;
RT "Cloning and expression of Xenopus Lrp5 and Lrp6 genes."
RL Mech. Dev. 117:337-342(2002).
DR EMBL; AF508961; AAN09807.1; -.
DR HSSP; P01130; 1AJJ.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR002172; LDL receptor A.
DR InterPro; IPR000033; Ldl_receptor_rep.
DR Pfam; PF00005; EGF; 4.
DR Pfam; PF00057; Ldl_recept_a; 3.
DR Pfam; PF00058; Ldl_recept_b; 19.
DR PRINTS; PR00261; LDLRECEPTOR.
DR SMART; SM00181; EGF; 4.
DR SMART; SM00192; LDLa; 3.
DR SMART; SM00135; LY; 20.
DR PROSITE; PS01209; LDLRA_1; 3.
DR PROSITE; PS00068; LDLRA_2; 3.
DR Lipoprotein; Receptor.
KW SEQUENCE 1613 AA; 180602 MW; BD93F09273DDDL77 CRC64;

Query Match		69.6%; Score 6084; DB 2; Length 1613;
Best Local Similarity		68.7%; Pred. No. 0;
Matches 1109; Conservative 206; Mismatches 280; Indels 20; Gaps 11;		
Qy	16	LLLLLALCGCPAPAAAPLLLFANRRDVRVLVADGVKLESTIVVSGLEDAADVDFQFSKG 75
Db	4	LLRSLLMCSLCLRASPLLIYANRRDLRLVDTAGMKGNSIVVSGLEDAADVDFVSRG 63
Qy	76	AVYWTDSBEAIAKTYLNQTAAYQNVVIVSGLVSPDGLACDWGKGLYWTDSNRIEVA 135
Db	64	LIYWSDVSEBAIKRDNKTKSS-QDVVIGLVSPDGLACDWLCKGLYWTDSNRIEVS 122
Qy	136	NLNGTSRKVLFWQDLQDPRATLAPAHGYMYWTWGETPRIERAGMGDSIRKIIIVDSDIY 195
Db	123	NLDGSLRKVIFWQDLQDPRATLAPAHGYMYWTWGEVPIKRIERAGMGDSVRSVIVDSDIY 182
Qy	196	WPNGLTLDLEQKLYWADAKLSFTHRANLDGSPKQVVEGSLTHPPFALTLLSGDLYWTDW 255
Db	183	WPNGLTLDYDEQKLYWADAKLSFTHKANMOGSHRQTVKGSALHPPFALTLEGDIIYWTDW 242
Qy	256	QTRSIHACNKRGTGKRKEILSALYSYPMIDQVLSQERQFFHTRCEEDNGGCSHCLLSPS 315
Db	243	TTRSILACNKYTGDRREVDTFISPMIDHVFSLRQPNATNPCAAHNGGCSHCLLSMPM 302
Qy	316	EPFYTCACPTGVQLODNGRTCKAGAEVLLARSTDLRRISLDTDPDTHVLVQDDIRHA 375
Db	303	EPFYQCACTGVRLMEGKKCMHGATELLALLVRRTDLRRISLDTDPDTHVLPLDDIRHT 362
Qy	376	IAIDYDPLEGVYVWTDDEVRAIRAYLDGSAQTLVNTINDPQIAVDWVARNLYWTDI 435
Db	363	IAIDYDVEGIYVWTDDEVRAIRAYTDQDSSQYIVTSQVAHPDQIAVDWVARNLYWTDI 422
Qy	436	GTDRIVTRLNGTSRKILVSEDLDEPRAIALHPVWGLMYWTDGSENPKECANLDGQERR 495
Db	423	GTDRIVTRLNGTWKRLISDLDEPRAIVLDPINGYMYWTDWGEIPKIECAAMDGSDRI 482
Qy	496	VLVNASIGWPNGLALDLOEGLYWGDAKTDEIVNVDGTERLTLEDKLPHEGFTLLG 555
Db	483	ILVNTSLGWPNGLALDTEGKIYWGDAKTDEIVNAIDGTRRVLVEDKLPHEGFTLLG 542
Qy	556	DFIYWTWQRRSIRERVHVKASRDVVIDQLPDLMLKAVNVAKVVTNPNPCADNRGGCSHL 615
Db	543	DYVVTWQRRSIRERVHKTGEREVIDQLPDLMLKATNTHKLSYTGCAENNGGCSHL 602
Qy	616	CFFTPHATRCGCPGLELLSMDKTCIVPEAFVFTSRAAIIHRISELTNNNDVAIPLTGK 675
Db	603	CLYRPOGPRCACPIGLELLINDMKTIVPEAFLLFSRRADIRRISELTNTVAIPLTGK 662
Qy	676	EASALDFDVSNHLYWTDSVLSKTIISRAFMNGSSVEHVVEGLDYPEGMAVDMGKLYWA 735
Db	663	EASALDFDVTDNRITYWTDVLSKTIISRAFMNGSALRHVVQGLDYPEGMAVDMGKLYWA 722
Qy	736	DTGTRNIEVARLDGQFVOLVYWRDLNPRSLALDPTKGIYIYWTWEGGKPRIVRAFMDCGN 795
Db	723	DTGTRNIEVSKLDGQHQILVWKOLDSRALALDPAEGFMYTWEGGKPKIDRTAMDGG 782
Qy	796	CMTLVKVRANDLTIDYADQRLYWTDLDTNMISSNMLGQERVVVIADDLPHPPGLTQYS 855
Db	783	RITLVDPVGRANGLTIDYAEERLYWTDLTDLIESNNMLGDRVVVIADDLPHPPGLTQY 842
Qy	856	DYIYWTWNHLSIRADKTSGRNRTLQGHLDVYMDILVHSSRODGLNDCWHNNGCGQ 915
Db	843	DYIYWTWSSRSIRANKTSQNRIMIQNDLDYVMDILVHSSRQAGWNECASSNGHCH 902
Qy	916	LCLALIP-GGHRGCGASHYTLDFSSRNCSPPFTFLFSQKSAISRMIPDDQHSPLIPLH 974
Db	903	LCLATPISGYTCGPVHSLTNDKTCAPSSFLFSQKNAINRMVIDGQSPDILPIH 962
Qy	975	GLRNVKAIYDPLKFIYVWQGRONIKRAKDDGTQFP-VLTSLSQGNPDQPHDLSDI 1033
Db	963	NLRNVRAIDYDPLEKLYWIDSRQNRIRRAQEDSGSMTIVASTIPNQNMQPYDLSIDI 1022
Qy	1034	YSRTLFWTCATNTINVHLSGEAMGVVLRGDRDKPRAIVVNAERGVLFTNMQDRAKI 1093
Db	1023	YSRLIYWCATNLIINTRLDGRAIGVVLKGEQRPRAILVNPERGMYFTNLQERSPKI 1082
Qy	1094	ERAAALDGEREVLFTTGLIRPVALVDNTLGLKFWVDADLKRIBSCDLSGANRUTLEDAN 1153
Db	1083	ERAAALDGEREVLFTTGLSKPVALALDNKMGKLPWADSLRRRIESSDLSGGHRIVLESDN 1142
Qy	1154	IVQPLGLTILGKHLIYWDROOQMIERVEKTTGDKRTIQRVAHLTGIHAEVSVLEBFS 1213
Db	1143	VLOPVGLTIFGNLYIWDROOQMIERIEKTRGVGRTKIOARIPLLTDIHAVTELIMDEYK 1202
Qy	1214	AHPCARDNGCSHICIAKGDGTPRCSPVHLVILQ-NULLTCGEBPTCSPDQFACATGID 1272
Db	1203	EHPCSQNGGCSHICIVKGDGTRCSPHLVLLQEBELSCGEBPTCSPHLFTCTGEVD 1262
Qy	1273	CIPGAMRCDFPEBCDDQSBEGCPVCSAAPPCARGQCVDLRLRCDEADQCQRSDAEC 1332
Db	1263	CIPAEWRCDFTECVDSHSDQNCPCMSDMQVQCTSGGICDSSLRCNGEDNCQDKSDEKNC 1322
Qy	1333	DAICLPNQPRCAGQCVLIKQOCDSPDCIDGSDLMCEITKPPSDSDSPAHSAGIPVIG 1392
Db	1323	NEVCAPDQFHCQSGQCIGKRRCDLSPDCSDSDSDEQACYPTEEP---PPPSNTNTIGSIIG 1379
Qy	1393	IILSLFVWGGVYFVQVVCORVYAGANGPPPEY-VSGTPEHVLNFIAPGSGHQGPFTGI 1451
Db	1380	VILTFVVGGVYFICQVLCRRMKGDGETWINDVYVHGVSVPLAYVPHPSLSLGSLEPM 1439
Qy	1452	ACGKSMSSVSLMGRCGVPLYDRNHVTGASSSSSSSTKATLYPPIILNPPSPATDPSLY 1511
Db	1440	SRGKSVIGSLISIMAGSGGP-YDRAHVTGASSSSSSSTKGTFFYPPPIILNPPSPATERSHY 1498
Qy	1512	NMDWFYSNTPATAR--PYRPIIRGMAPPTTPCSTDCVSDYS-----ASRWKASKYLL 1564
Db	1499	TMEFGYSNSPSTRSYSPYRPHFAPPTTPCSTDCVSDYTPSHRLTSAAAAGKYS 1558
Qy	1565	DLNDSDPYPPPPPHSQYLSAE---DSCPPSPATERSY-FHLPPPPPPSPCTDSS 1615
Db	1559	DLNVDSBPVPPPPPHSQYLSAEENYESCPSPTERSYSHHLYPPPPSPCTDSS 1613
RESULT 8		
Q95V09		PRELIMINARY; PRT; 1678 AA.
ID	Q95V09	
AC	Q95V09	
DT	01-DEC-2001	(TrEMBLrel. 19, Created)
DT	01-DEC-2001	(TrEMBLrel. 19, Last sequence update)
DT	01-MAR-2004	(TrEMBLrel. 26, Last annotation update)
DE	Arrow.	
GN	Name=arr;	
OS	Drosophila melanogaster (Fruit fly).	
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;	
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;	
OC	Ephydroidea; Drosophilidae; Drosophila.	
OX	NCBI_TaxID=7227;	
EN	[1]	
RP	SEQUENCE FROM N.A.	
RA	Wehrli M., Dougan S.T., Caldwell K., O'Keefe L., Schwartz S.,	
RA	Vaizel-Ohayon D., Schejter E., Tomlinson A., DiNardo S.;	
RT	"arrow encodes an Ldl receptor-related protein essential for Wingless	
RT	signaling in Drosophila."	
RL	Nature 0:0-0(2000).	
DR	EMBL; AY005815; AAF91072.1; --	
DR	HSSP; P01130; IN7D.	
DR	Flybase; FBgn0000119; arr.	
DR	GO; GO:0016055; P:Wnt receptor signaling pathway; IMP.	
DR	InterPro; IPR000152; Asx hydroxyl_s.	
DR	InterPro; IPR006209; EGF_like.	
DR	InterPro; IPR006210; IEGF.	
DR	InterPro; IPR002172; Ldl_receptor_A.	
DR	InterPro; IPR000033; Ldl_receptor_rep.	
DR	Pfam; PF00008; EGF; 2.	
DR	Pfam; PF00057; Ldl_recept_a; 3.	
DR	Pfam; PF00058; Ldl_recept_b; 16.	

Db 393 SANTCANGQEMFIVQRTQISKISLSDSPDYTFIFPLGKVKYKAIADYDPEEHIYNSD 452
Qy 392 DEVRAIRAYLDGSAQTLNTEINDPGIAVDWVARNLYWTDGTDRIEVTRLNGTSRK 451
Db 453 VETYIKRAHADGTGTVTFVSEVRHPDGLADWLARNLYWTDVIXDRIEVCRIDGTARK 512
Qy 452 ILVSDLEDEPRALALHPVWGLMYWTDWGE-NPKIECANLDGOERRVLYNASIGWENGLAL 510
Db 513 VLIYEHLEPREAIAVAPSLGWMFWSDMNERPKVERASLDGSEVRVLYVSENLGWPNGLAL 572
Qy 511 DLQEGKLYWGDAKTDKIEVINVDGTRKRTLLLEDKLPHIFGFTLLGDFTYWDWQRSTIER 570
Db 573 DIEAKAIYWCDBKTDKIEVANWMDGSGRRVVISDNLKHLFGLSILDYLYWTDWQERSIDR 632
Qy 571 VHKVKA-SRDVLIIDQLPOLMGLKAVNAKVVTNTPCADRNGGCSHLCTFFTHATRCGCP 629
Db 633 AHKITGNRNVVVDQYPLDMLGLKVTRLREVRQNAACAVRNGGCSHLCLNRPDRVYCRCAI 692
Qy 630 GLELLSDMKTICVPEAFIVFTSRAAIHRISL---ETNNNDVAIPLTGVKEASALDPOYSN 686
Db 693 DYELANDKRTCVVPAFLFRQEHIGRISIEYNEGNHDERIPKQVDRDAHALDVSAE 752
Qy 687 NHIYWDVSLKTIISRAFNMGSSVEHVFEGLDYPEGMAVDMWGNKLYWADTGTNRIEVAR 746
Db 753 RRIYWDQSKICIPRAFNGSVQRIVDSGLIGPDGIAVDMLANNIYWSDAEARIEVAR 812
Qy 747 LDGQFQVLRDLNPRSLALDPTKGYIYTWGCKGPRIVRAFMGNGNCTMLVDKVGRA 806
Db 813 LDGSSRRVLLMKGVEPRSLVLEPRRGYMYTE-SPTDSIRRAAMDGDQLTIVAGANHA 871
Qy 807 NDLTIDYADQLYV-TDLDTNMISNNMGOER-VVIADDLPHPRGLTQYSDYLYWTDWN 864
Db 872 AGLTFDQETRLRYWATQSRPAKISADWDGKRRQLLVGSDMDPEYAVSLYQDYVWSDWN 931
Qy 865 LHSIERADKTSGRNRTLQGHLDVMDILVPHSSRRQDGLNDCMNNNGCGOLCLAIPO-- 922
Db 932 TGDIERVHKTTQONSLRVHSGMTVYITSLVFNDRKQGTGVNPKVNGGCSHLCLAQPCR 991
Qy 923 GHRCCASHYTLDPSSRNCSPPTFLFESQSAISRMTPDQHSPLDILPLHGLRNKVAI 982
Db 992 GMTACAPHYQAKDGVGSCIPFRNYIFSRQNCFRLLPNTDPCNPLPVSG-KNIRAV 1050
Qy 983 DYDPLDKFIYWDGR-QNIKRAKDDGTQFFVLTSLSQONPDROPDHLSDIYSRTLPWT 1041
Db 1051 DYDPIITHIYIEGRSHISKSLANGTKVSLAN-----SGQPPDLAIDIGELLFWT 1103
Qy 1042 CEATNTINVHLSGEMGVLRGDRKPRIVNAERGVLVFTNMQDRAAKIERAALDGT 1101
Db 1104 CSQNSINVTSLFSGVGVITDGSSEKPRNIAVHAMKRLFWTDVGSQA-IIRARVDGN 1162
Qy 1102 BREVLTFTGLIRPVALVVDNTLGLFWVDADLKRIESCDLSGANRLTLEDANIVQPLGLT 1161
Db 1163 ER-VELAYKLGVTALADQSDMIY--AHGKRIDAIDINGKNKTLVSMHISQVINIA 1219
Qy 1162 ILGKHLWIDRQQQMIERVEXTGDKRTIRIQGRAVHTGTHAVEEVSLEEPSAHPCARDN 1221
Db 1220 ALGSGFVWLD-DKTGVERI-TVNGERSAEQLRPLQITDIRAVWTPDPKVLNHTCHSR 1277
Qy 1222 GGCSHICIAKGDTGTPR-----CSCPVLVLLQNLITCGEPPTSCSPQACATG-----E 1270
Db 1278 TKSHICIASGEGARTDRVDCPCPKHMLLMDKENCGAFACGPDHFTCAAPVSGISDVN 1337
Qy 1271 IDCIPGAWRCDGPFCDOSDEECPCVCSAAQFPFCAGQCVDLRLRCGDEADCODRDEA 1330
Db 1338 KDCIPASWRCDQKDCPDKSEVCGCTCRADQFCQSGECIDKSLVCDGTTNCANGHDEA 1397
Qy 1331 DCDAIC-LPNQFRCASGO-CVLIKQCCDSPFDCIDGSD---LMCEITKPPSDPSAHS 1385
Db 1398 DC---CKRPGFQCPINKLCISAALLCDGWENCADGADDESSDICLQRRMAPATDKRAFI 1454
Qy 1386 AIGPVGIGLILSLFWGVGYVFCQVRCVQYAGANGPPHPHEVSGTBPVPLNFIAPGSQH 1445
Db 1455 LIGATMITIFSI-----VYLL---QFCRTRIGKSRTEPKDQADTDLPSPTL-----SKS 1501

RESULT 11

Q7FV65 PRELIMINARY; PRT; 1698 AA.
ID Q7FV65 AC Q7FV65;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ENSANGP0000020132 (Fragment).
GN Name=ENSANGG00000017643;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (Apr-2003) to the EMBL/GenBank/DBJ databases.
CC Sub-CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAA801008986; EAA00402.2; -.
DR HSSP; Q07954; 1D2L.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR000152; Aax_hydroxyl_S.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR002172; LDL_receptor_A.
DR InterPro; IPR000033; LDL_receptor_rep.
DR InterPro; IPR001005; Myb_DNA_binding.
DR Pfam; PF00008; EGF_1.
DR Pfam; PF00057; Ldl_recept_a; 3.
DR Pfam; PF00058; Ldl_recept_b; 17.
DR PRINTS; PR00261; LDLRECEPTOR.
DR PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_1.
DR PROSITE; PS01209; LDLRA_1; 2.
DR PROSITE; PS00068; LDLRA_2; 1.
DR PROSITE; PS00037; MYB_1; UNKNOWN_1.
FT NON TER.
SQ SEQUENCE 1698 AA; 187081 MW; 2498484BD5044E93 CRC64;

Query Match 37.3%; Score 3262.5; DB 2; Length 1698;
Best Local Similarity 40.7%; Pred. No. 5.8e-196;
Matches 696; Conservative 277; Mismatches 580; Indels 157; Gaps 44;

Qy 27 APAASP---LLLFANRRDVRV-----DAGGVKLESTIVSGLEDAADVDFSGAV 77
Db 21 ATAASTSTILLFTTYDYDLRVNLVSVQPDNTALTYMDVLQDLQNEASAMDFHARGLV 80
Qy 78 YWTDVSEALIKQTYLNTGAAVQ--NVVISGLVSPDGLADWVGKLYTWDSETNRIEVA 135
Db 81 CWIENTLEVIOCCGTNGTGLSVTSKTTVTITDGLDKPEGLADWYTDKLYWTDGESNRIEVA 140
Qy 136 NL-----NGTSRKVLFWODLQOPRAIALDPAHGYMYWTDWG 171
Db 141 QLEPFGGGGAGATGGGGGGIGGSKQPNARLQKVLIWSDLPQRAIALVPARRYMIWTDWG 200

RP CONCEPTUAL TRANSLATION.
 RA Blatter M.-C.;
 RL Unpublished observations (MAR-2004).
 RN [3]
 RP SEQUENCE OF 608-1905 FROM N.A., AND TISSUE SPECIFICITY.
 RC STRAIN=Sprague-Dawley; TISSUE=Brain;
 RX MEDLINE=98360089; PubMed=9693030; DOI=10.1006/geno.1998.5341;
 RA Nakayama M., Nakajima D., Nagase T., Nomura N., Seki N., Ohara O.;
 RT "Identification of high-molecular-weight proteins with multiple EGF-like motifs by motif-trap screening";
 RL Genomics 51:27-34(1998).
 CC -!- FUNCTION: Potential cell surface endocytic receptor, which binds and internalizes extracellular ligands for degradation by lysosomes.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
 CC -!- TISSUE SPECIFICITY: Expressed in different regions of the brain, mainly in the olfactory bulb, at lower level in the cerebral cortex and hippocampus.
 CC -!- SIMILARITY: Belongs to the LDLR family.
 CC -!- SIMILARITY: Contains 3 EGF-like domains.
 CC -!- SIMILARITY: Contains 8 LDL-receptor class A domains.
 CC -!- SIMILARITY: Contains 15 LDL-receptor class B domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AB011533; BAA88688.1; -
 DR RGD; 619731; Ltp4.
 DR InterPro; IPR006209; EGF like.
 DR InterPro; IPR006210; IEGF.
 DR InterPro; IPR000033; Ldl_receptor_rep.
 DR Pfam; PF00008; EGF; 3.
 DR Pfam; PF00058; Ldl_recept_b; 16.
 DR SMART; SM00181; EGF; 3.
 DR SMART; SM00135; LY; 17.
 DR PROSITE; PS00010; ASK_HYDROXYL; 1.
 DR PROSITE; PS00022; EGF_1; FALSE_NEG.
 DR PROSITE; PS01186; EGF_2; 3.
 DR PROSITE; PS00026; EGF_3; FALSE_NEG.
 DR PROSITE; PS01187; EGF_CA; 1.
 DR PROSITE; PS01209; LDLRA_1; 8.
 DR PROSITE; PS00068; LDLRA_2; 8.
 KW Calcium-binding; EGF-like domain; Endocytosis; Glycoprotein; Receptor;
 KW Repeat; Signal; Transmembrane.
 FT SIGNAL 1 20 Potential.
 FT CHAIN 21 1905 Low-density lipoprotein receptor-related protein 4.
 FT FT Extracellular (Potential).
 FT FT Potential.
 FT DOMAIN 21 1723 Cytoplasmic (Potential).
 FT TRANSMEM 1724 1746 LDL-receptor class A 1.
 FT DOMAIN 1747 1905 LDL-receptor class A 2.
 FT DOMAIN 26 67 LDL-receptor class A 3.
 FT DOMAIN 70 106 LDL-receptor class A 4.
 FT DOMAIN 109 144 LDL-receptor class A 5.
 FT DOMAIN 147 183 LDL-receptor class A 6.
 FT DOMAIN 190 226 LDL-receptor class A 7.
 FT DOMAIN 230 266 LDL-receptor class A 8.
 FT DOMAIN 269 305 LDL-receptor class A 9.
 FT DOMAIN 311 350 EGF-like 1 (atypical).
 FT DOMAIN 354 394 EGF-like 2, calcium-binding.
 FT DOMAIN 395 434 EGF-like 3.
 FT DOMAIN 480 521 LDL-receptor class B 1.
 FT DOMAIN 523 564 LDL-receptor class B 2.
 FT DOMAIN 566 608 LDL-receptor class B 3.
 FT DOMAIN 610 651 LDL-receptor class B 4.
 FT DOMAIN 653 693 LDL-receptor class B 5.
 FT DOMAIN 698 737 EGF-like 3.
 FT DOMAIN 785 826 LDL-receptor class B 6.
 FT DOMAIN 828 869 LDL-receptor class B 7.

FT	DOMAIN	871	913	LDL-receptor class B 8.
FT	DOMAIN	915	956	LDL-receptor class B 9.
FT	DOMAIN	957	997	LDL-receptor class B 10.
FT	DOMAIN	1093	1134	LDL-receptor class B 11.
FT	DOMAIN	1136	1177	LDL-receptor class B 12.
FT	DOMAIN	1179	1221	LDL-receptor class B 13.
FT	DOMAIN	1223	1263	LDL-receptor class B 14.
FT	DOMAIN	1264	1305	LDL-receptor class B 15.
FT	DOMAIN	1397	1438	LDL-receptor class B 16.
FT	DOMAIN	1440	1481	LDL-receptor class B 17.
FT	DOMAIN	1483	1525	LDL-receptor class B 18.
FT	DOMAIN	1527	1572	LDL-receptor class B 19.
FT	DISULFID	27	44	By similarity.
FT	DISULFID	34	57	By similarity.
FT	DISULFID	51	66	By similarity.
FT	DISULFID	71	83	By similarity.
FT	DISULFID	78	96	By similarity.
FT	DISULFID	90	105	By similarity.
FT	DISULFID	110	122	By similarity.
FT	DISULFID	117	135	By similarity.
FT	DISULFID	129	143	By similarity.
FT	DISULFID	148	160	By similarity.
FT	DISULFID	155	173	By similarity.
FT	DISULFID	167	182	By similarity.
FT	DISULFID	191	203	By similarity.
FT	DISULFID	198	216	By similarity.
FT	DISULFID	210	225	By similarity.
FT	DISULFID	231	243	By similarity.
FT	DISULFID	238	256	By similarity.
FT	DISULFID	250	265	By similarity.
FT	DISULFID	270	282	By similarity.
FT	DISULFID	277	295	By similarity.
FT	DISULFID	289	304	By similarity.
FT	DISULFID	312	324	By similarity.
FT	DISULFID	319	337	By similarity.
FT	DISULFID	331	349	By similarity.
FT	DISULFID	358	369	By similarity.
FT	DISULFID	365	378	By similarity.
FT	DISULFID	380	393	By similarity.
FT	DISULFID	399	409	By similarity.
FT	DISULFID	405	418	By similarity.
FT	DISULFID	420	433	By similarity.
FT	DISULFID	702	713	By similarity.
FT	DISULFID	709	722	By similarity.
FT	DISULFID	724	736	By similarity.
FT	CARBOHYD	264	264	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	498	498	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	719	719	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	901	901	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	1077	1077	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	1415	1415	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	1467	1467	N-linked (GlcNAc. . .) (Potential).
SO	SEQUENCE	1905 AA; 211878 MW; 9562AS729D69E29A	CRC64;	

Query Match 30.7%; Score 2687; DB 1; Length 1905;
 Best Local Similarity 43.0%; Pred. No. 1,1e-159;
 Matches 539; Conservative 222; Mismatches 443; Indels 50; Gaps 21;

Qy	33	PLLLFANRRDVLVDAGGVKLESTIVVSGLEDAADVQFQSKGAVYWTVDVSEEAIKQTYL	92
Db	440	PVLLFANRDIRQVLPH--RSEYTLNLENAIADPFHRRRLVFWSDVTLDRILRANL	497
Qy	93	NOTGAAQNVVVISGLVSPDGLCDWVGKLYWTDSETNRLEVANLNGTSRKVLFWODLDQ	152
Db	498	N--GSNVEEVVSTGLSPGGLAVDWHDKLYWTDGTSRLEVANDGAHRKVLWQSLQK	555
Qy	153	PRAIALDPAGHYMYTWDGTPPRIERACMDGSTRKIIIVSDIYWPNGLITDLEOKLYWA	212
Db	556	PRAIALHPMEGTIYWDGNTGTPRIEASSMDGSGRRIIADTHLFWPNGLITDYAGRRYVW	615
Qy	213	DAKLSFTHRANLDGSRFKQKVVESGLTHPPFALTLSGDTLYWTDQWTSRSHACNKTGGK	272
Db	616	DAKHVIERANLDGSHRKAVISQGLPHFPALTVPEDSLYWTDWHTKTSINSANKFTGKQ	675

```
QY 273 EILSALYSPMDIQVLQSOEROPFFHTRCEEDNGGCHLCLLSSEBFFYTACPTGVQLQDN 332
Db 676 IIRNKLHPMDIHTLHPQRPAGKRCNGDNGGCHLCL--PSGQNYTACPTGFR-KIN 732
QY 333 GRTCKAGAEVLLARTRIRISLTDPTDVLQVDDIRHAIADVDPLRGYVYVTD 392
Db 733 SHACAQLDKFLFARRMDIRISFDTEDSDVPLADVRSAVALDMSRDDHVIWTDV 792
QY 393 EVRAIRRAYLDGSGAQTLVNTEINDPDGIADVARNLYWTDGTDRIVETRLNGTSRK 452
Db 793 STDTISRAKWDGTGQKVVVDTSLESPAGLAIDWVTKLYWTDAGTDRIEVANTDGSMTV 852
QY 453 LVSEDLDEPRATALHPVCLWYTWNGENPKTECANLQGERVLVNASLWPNGLADL 512
Db 853 LIWENLDRPRIVBFGMGYWTWDGASPKTIERAGMDASNRQVITISNLTPNGLAID 912
QY 513 QEGKLYWDGAKTDKLEVINVDGTEKRTILLEDKLPHIFGFTLLGDFIYWDQRRRIERVH 572
Db 913 GSORLYWADAGMKTIEFAGLDGSKKVLIGSQLPHFGTLGQRIYIYWDQTKSIQAD 972
QY 573 KVKA-SRVIDIQLPDLMLKAVNVAKVGNTPCADRNGGCHLCLFFTPHAT--RCGCP 629
Db 973 RLTLGLDRETLQENLENLMDIHVFHRQRPVITPCAVENGCGCHLCLRSPSPGFCCTPT 1032
QY 630 GLELLDMKTCIV-PEAFIVETSRAAIHRI SLETN--NDVAIPLT-GYKEASALDFVSN 686
Db 1033 GINLLDGKTCPGMNSFLIFARRIDVRVSLDIPYFADVVVPINMTWKNTIAIGVDPLE 1092
QY 687 NHIYWDVSLKTIISRAFNGSSVEHVFEFGDYPGMAVDMWGNKLYWADGTNRIEVAR 746
Db 1093 GKVYWSDSLTHRISSRSLDGSQHEDIITGLQTTDGLAVDAIGRKVYWDGTGTRIEVGN 1152
QY 747 LDGQFQVLVWRDLNPRSLALDPTKGIYIYTWGCKPRIIVRAFMDGNCMTLV-DKVR 805
Db 1153 LDGSMKRLVWQNLDSRAIYLHEMGFYMTDGENAKLERSGMDGSDRTVLINNLGW 1212
QY 806 ANDLTDVADQRYWTDLTWNIESNMLQGERVVIADLPHPLGLTOYSDYIYWTWN 865
Db 1213 PNLTVDTKTSQLLWADHAETRIEVLADGANRHTLVSPQHPGLTLLSDYIYWTDMQT 1272
QY 866 HSIERADTKSGRNTLIQGLHLDVMDILVFFHSRQDGLNDQMNHNGCGQLCLAIPEGHR 925
Db 1273 RSIHRADKSTGNSVILVRESNLPGLMDIOAVDRAQPLGFKNGSRNGGCHLCLRPSGFS 1332
QY 926 CGCASHYITLDPSSRNCSP-PTFLLFPSQSAISRMI PDDQHSPLDILPLHGLRNKAIDY 984
Db 1333 CACPTGIQLKGDGKTCDPSPETIYLLFSSRGSIRISLTDHTDTHVVPVGLNVLSDY 1392
QY 985 DPLDKFIYWDGRON-IKRAKDDGTQPPVLTSLSQGNPDPROPHDLSIDISRTLFWTCE 1043
Db 1393 DSDVGKYYTDFVLDVIRADLNSN--METVIGHGL---KITDGLADVWVARNLYWTD 1447
QY 1044 ATNTINVHLSGEAMGVVLGRDKPRAIVVNAERGYLEFTNMQDRAAKIERAALDGTFR 1103
Db 1448 GRNTIEASRLDGSCKVLINNSLDEPRAIAVPRKGLFWTDW-GHIAKIERANLDGSR 1506
QY 1104 EVLFTTGLIRVALVWNTLCKLFWADLAKRIESCDSGANRLTLEDANIVQPLGLTIL 1163
Db 1507 KVLINADLGWPNGLTLDYDTRIIYVWDADLRIESADLNGKLRQVLV-SHVSHPFALTQ 1565
QY 1164 GKHLWIDRQOQMIERVEKTTGDKRTIQGRVAHLTGTHAVEEVSLE-EPSAHPCARDNG 1222
Db 1566 DRWIYWTDMQTKSORVDKYSGRNKETV--LANVEGLMDIIVVSPQRTGTNACGVNG 1622
QY 1223 GCSHICIAKGDTPRCSCVHLVLLQNLKTCGPPTCSPDQFACATGEIDCIPG 1276
Db 1623 GCSHLCFARASDFV-CACP-----DEPDSHPGCS-----LVPG 1653

PRELIMINARY; PRT; 1905 AA.
```

RESULT 14

Q76LU2

ID Q76LU2

```
AC Q76LU2;
DT 05-JUL-2004 (TRENBLrel. 27, Created)
DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
DE LDL receptor-related protein 13.
GN Name=LRP13;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Tian Q., Suzuki T., Okano A., Usuda N.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB073317; BAD18061.1; -.
DR HSSP; P01130; 1AJJ.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR InterPro; IPR000152; Asx_hydroxyl_S.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR002172; LDL_receptor_A.
DR InterPro; IPR000033; Ldl_receptor_rep.
DR Pfam; PF00008; EGF; 4.
DR Pfam; PF07645; EGF_CA; 1.
DR Pfam; PF00057; Ldl_recept_a; 8.
DR Pfam; PF00058; Ldl_recept_b; 19.
DR PRINTS; PR00261; LDLRECEPTOR.
DR SMART; SM00181; EGF; 7.
DR SMART; SM00179; EGF_CA; 2.
DR SMART; SM00192; LDLA; 8.
DR SMART; SM00135; LY; 20.
DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR PROSITE; PS01186; EGF_2; 3.
DR PROSITE; PS01187; EGF_CA; UNKNOWN 1.
DR PROSITE; PS01209; LDLA_1; 7.
DR PROSITE; PS00068; LDLRA_2; 8.
KW EGF-like domain; Receptor.
SQ SEQUENCE 1905 AA; 211928 MW; 4A9ACB7D00ED54FE CRC64;

Query Match 30.7%; Score 2687; DB 2; Length 1905;
Best Local Similarity 43.0%; Pred. No. 1.1e-159;
Matches 539; Conservative 222; Mismatches 443; Indels 50; Gaps 21;

QY 33 PLLLFANRRDRLVLDAGGVKLESTIVVSGLEDAADVDFQFSKGAVYVTDVSEEAIKOTYL 92
Db 440 PVLFFANRIDIRQVLPH--RSEYTLNLLNLENAIALDFHRRRLVFWSDVTLDRILRANL 497
QY 93 NOTGAAVQNVVIGSLVSPDGLACDWGKKLYWTDSETNRIEVANLNGTSRKVLFWQLDQ 152
Db 498 N--GSNVEEVVSTGLESFPGGLAVDWVHDKLYWTDGTSRLEVANLDAHRKVLWQSLK 555
QY 153 PRAIALDPAGHYVWTDWGTPTRIERAGMDGSTRKIIVDSDIYWPNGLTITDLERQKLYWA 212
Db 556 PRAIALHPMGTYITWTDWGTPTRIERAGMDGSTRKIIVDSDIYWPNGLTITDIYAGRIWV 615
QY 213 DAKLSFTHIRANLDGSRQKQVVEGSLTHPFAITLSGDTLYWTDWQTRSIHACNKTGGKRK 272
Db 616 DAKHHVIERANLDGSHRKAVISQGLPHFPFAITVFEDSLYWTDWHTKSINSANKFTGNQE 675
QY 273 EILSALYSPMDIQVLQSOEROPFFHTRCEEDNGGCHLCLLSSEBFFYTACPTGVQLQDN 332
Db 676 IIRNKLHPMDIHTLHPQRPAGKRCNGDNGGCHLCL--PSGQNYTACPTGFR-KIN 732
QY 333 GRTCKAGAEVLLARTRIRISLTDPTDVLQVDDIRHAIADVDPLRGYVYVTD 392
Db 733 SHACAQLDKFLFARRMDIRISFDTEDSDVPLADVRSAVALDMSRDDHVIWTDV 792
QY 393 EVRAIRRAYLDGSGAQTLVNTEINDPDGIADVARNLYWTDGTDRIVETRLNGTSRK 452
Db 793 STDTISRAKWDGTGQKVVVDTSLESPAGLAIDWVTKLYWTDAGTDRIEVANTDGSMTV 852
```


DR SMART; SM00179; EGF CA; 1.
DR SMART; SM00135; LV; 20.
DR PROSITE; PS00010; ASX HYDROXYL; 1.
DR PROSITE; PS00022; EGF_1; FALSE_NEG.
DR PROSITE; PS01186; EGF 2; 3.
DR PROSITE; PS00026; EGF 3; FALSE_NEG.
DR PROSITE; PS01187; EGF CA; 1.
DR PROSITE; PS01209; LDLRA_1; 8.
DR PROSITE; PS00068; LDLRA_2; 8.
KW Calcium-binding; EGF-like domain; Endocytosis; Glycoprotein; Receptor;
Repeat; Signal; Transmembrane.
FT SIGNAL 1 20
FT CHAIN 21 1950
FT DOMAIN 21 1768
FT TRANSMEM 1769 1791
FT DOMAIN 1792 1950
FT DOMAIN 71 112
FT DOMAIN 115 151
FT DOMAIN 154 189
FT DOMAIN 192 228
FT DOMAIN 235 271
FT DOMAIN 275 311
FT DOMAIN 314 350
FT DOMAIN 356 395
FT DOMAIN 399 439
FT DOMAIN 440 479
FT DOMAIN 525 566
FT DOMAIN 568 609
FT DOMAIN 611 653
FT DOMAIN 655 696
FT DOMAIN 698 738
FT DOMAIN 743 782
FT DOMAIN 830 871
FT DOMAIN 873 914
FT DOMAIN 916 958
FT DOMAIN 960 1001
FT DOMAIN 1002 1042
FT DOMAIN 1138 1179
FT DOMAIN 1181 1222
FT DOMAIN 1224 1266
FT DOMAIN 1268 1308
FT DOMAIN 1309 1350
FT DOMAIN 1442 1483
FT DOMAIN 1485 1526
FT DOMAIN 1528 1570
FT DOMAIN 1572 1617
FT SITE 1611 1814
FT DISULFID 72 89
FT DISULFID 79 102
FT DISULFID 96 111
FT DISULFID 116 128
FT DISULFID 123 141
FT DISULFID 135 150
FT DISULFID 155 167
FT DISULFID 162 180
FT DISULFID 174 188
FT DISULFID 193 205
FT DISULFID 200 218
FT DISULFID 212 227
FT DISULFID 236 248
FT DISULFID 243 261
FT DISULFID 255 270
FT DISULFID 276 288
FT DISULFID 283 301
FT DISULFID 295 310
FT DISULFID 315 327
FT DISULFID 322 340
FT DISULFID 334 349
FT DISULFID 357 369
FT DISULFID 364 382
FT DISULFID 376 394
FT DISULFID 403 414

FT DISULFID 410 423
FT DISULFID 425 438
FT DISULFID 444 454
FT DISULFID 450 463
FT DISULFID 465 478
FT DISULFID 747 758
FT DISULFID 754 767
FT DISULFID 769 781
FT CARBOHYD 309 309
FT CARBOHYD 543 543
FT CARBOHYD 764 764
FT CARBOHYD 946 946
FT CARBOHYD 1122 1122
FT CARBOHYD 1460 1460
FT CARBOHYD 1512 1512
FT CONFLICT 1131 1131
FT CONFLICT 1599 1599
FT CONFLICT 1691 1691
FT CONFLICT 1907 1907
SQ SEQUENCE 1950 AA; D298624D70B2A287 CRC64;
Query Match 30.7%; Score 2686; DB 1; Length 1950;
Best Local Similarity 42.8%; Pred. No. 1.4e-159;
Matches 541; Conservative 229; Mismatches 447; Indels 46; Gaps 21;
QY 33 PLLLFANRRDRLVDAGVKLESTIVVSGLEDAADVDFQSKGAVYWDVSEBAIKQTYL 92
DB 485 PVLFFANRDIRQRLPH--RSEYTLLENLENAIADLDFHRRELVPFWSVDVTLRLANL 542
QY 93 NOTGAAQNVNVISGLVSPDGLACDWGKLYWTDSETRNRIEVANLNGTSRKVLWQDLQD 152
DB 543 N--GSNVEEVVSTGLESPGGLAVDWVHDKLYWTDSTGTSRIEVANLGDANRKLWQNLK 600
QY 153 PRAIALDPANGYMYWTDWGTPTPRIERAGMDGSTRKIIIVSDIYWPNGLTIDLEBQKLYWA 212
DB 601 PRAIALHPMEGTIYWDGNTPTRIEASSMDGSGRRIIADTHLFWPGLTIDYAGRRMYWV 660
QY 213 DAKLSFTHRLNDSFRQKVEGSLTHPPFALTLSGDTLYWTDWOTRTHACNKETGGKRK 272
DB 661 DAGHVIERANLDGSHRKAVISQGLPHFPFAITVFEDSLYWTDWHTKSINSKGTGNQE 720
QY 273 EILSALYSPMDIQVLQSERQFFHTRCEBNDGGCSHLCLLSPSEFFYTCAPTQVQLQDN 332
DB 721 IIRNKLHFPMDIHTLHPQRPAGKRNKRGDNGGCTHCL--PSQNYTCACPTGFR-KIS 777
QY 333 GRTCKAGAEVLLARFTDLRRISLDPDFTDIVLQVDDIRHAIDAIDYDPLEGVYVWTD 392
DB 778 SHACAQSLDKFLFARRMDIRISFDTELSDDVPIPLADVRSVALDWDSDRDHVVYWTDV 837
QY 393 EVRAIRRAYLDGGAOFLVNTENDDPGIADVWVARNLYWTDGTGTRIEVTRNGTGRKI 452
DB 838 STDTISRAKWDGTQGVVVDITSLSPAGLAIDWVTKLYWTDAGTDRIEVANTDGSWRTV 897
QY 453 LVSEDLDEPRAIALHPVMGLMYWTDNGENPKIECANLDGQERRVNLNAGLWPNGLALDL 512
DB 898 LIWENLDRPRDIIVPEPMGGYMYWTDGASPKIERAGMDASGRQVLISSNLTPNGLAIDY 957
QY 513 QEGKLYWGDAKTKIEVINVDGTRKRTLLDEKLPHPFGTLGDFIYWTWQRRSIRVH 572
DB 958 GSORLYWADAGMKTIIEPAGLDGSKRKVIGSQLPHPGLTYGERIYWTWQTKSIQASD 1017
QY 573 KVKA-SRDVIIDQLPDLMLGKAVNVAKVCTNPACDRNGCSHLCPFTPHAT--RCGCP 629
DB 1018 RLTLGURETLEQENLENLMDHVFHRRPPVSTFCAMENGCSHLCLSPNPSGFCSTCPT 1077
QY 630 GLELLSDMKTCIV-PEAFVFTSRAAHRISLETNN-NDVAIPTL-GVKEASALDFDVSN 686
DB 1078 GINLLSDGKTCSPQWNSFLIFARRIDIRMVSLDIPYFADVVPVINITMKNTIAIGVDPQE 1137
QY 687 NHIYWTDVSLKTIISRAFMNGSSVHVVEFGLDYPEGMVNDWGMKNLYWATGTINRIEVAR 746
DB 1138 GKVYWSDSLTHIRSRANLDGSHQEDIIITGLQTTDGLAVDAIGRKVYWTDTGTNRIEVGN 1197

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 17, 2005, 01:24:44 ; Search time 58 Seconds
(without alignments)
2679.139 Million cell updates/sec

Title: US-09-931-375A-2

Perfect score: 8740

Sequence: 1 MEAAPGPPWPLLLLLLLLL.....TERSYPFLPPPPSPCTDSS 1615

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_79:*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8736	100.0	1615	2 J03072	low density lipopr
2	6153.5	70.4	1613	2 J02072	low density lipopr
3	6097.5	69.8	1613	2 J02073	low density lipopr
4	2654	30.4	527	2 J03073	low density lipopr
5	2193	25.1	4544	1 S02392	alpha-2-macroglobu
6	2180	24.9	4543	1 A53102	alpha-2-macroglobu
7	2113	24.2	4545	1 S25111	alpha-2-macroglobu
8	1856	21.2	4660	2 T42737	gp330 protein prec
9	1458	16.7	4753	1 A47437	LDL-receptor-relat
10	1024	11.7	252	2 T46336	hypothetical prote
11	914.5	10.5	1847	2 T18308	probable vitellog
12	913.5	10.5	1984	2 T13171	probable vitellog
13	902	10.3	1650	2 S53457	dominant autoantig
14	824.5	9.4	1207	1 EGHU	epidermal growth f
15	805	9.2	1217	1 EGM5MG	epidermal growth f
16	794	9.1	2643	2 T29149	hypothetical prote
17	753	8.6	1133	1 EGRT	epidermal growth f
18	735	8.4	2215	2 T00348	LR11 protein - mou
19	663	7.6	2180	2 T29764	hypothetical prote
20	651.5	7.5	863	1 S51789	LDL receptor prec
21	648.5	7.4	873	1 QRRBVD	LDL receptor prec
22	648	7.4	873	1 A49729	LDL receptor prec
23	643	7.4	873	1 I48952	LDL receptor prec
24	634.5	7.3	1661	2 T31330	head-activator bin
25	617.5	7.1	837	1 A29512	LDL receptor prec
26	611.5	7.0	996	2 J02037	apolipoprotein E r
27	608	7.0	909	1 QRXL1	LDL receptor 1 pre
28	607.5	7.0	862	1 QRMSLD	LDL receptor precu
29	606	6.9	860	1 QRHULD	LDL receptor precu

ALIGNMENTS

RESULT 1

JE0372

low density lipoprotein receptor related protein - human

C:Species: Homo sapiens (man)

C:Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 09-Jul-2004

C:Accession: JE0372; JE0329

R:Kim, D.; Inagaki, Y.; Suzuki, T.; Ioka, R.X.; Yoshioka, S.Z.; Magoori, K.; Kang, M.; Cl

J. Biochem. 124, 1072-1076, 1998

A:Title: A new low density lipoprotein receptor related protein, LRP5, is expressed in he

A:Reference number: JE0372; MUID:99054722; PMID:9832610

A:Accession: JE0372

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1615 <KIM>

A:Cross-references: UNIPROT:075197; DDBJ:AB017498

R:Dong, Y.; Lathrop, W.; Weaver, D.; Qiu, Q.; Cini, J.; Bertolini, D.; Chen, D.

Biochem. Biophys. Res. Commun. 251, 784-790, 1998

A:Title: Molecular cloning and characterization of LR3, a novel LDL receptor family prote

A:Reference number: JE0329; MUID:99008902; PMID:9790987

A:Accession: JE0329

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1329, 'A', 1331-1615 <DON>

A:Cross-references: GB:AF077820; NID:G3831747; PIDN:AACT2791.1; PID:G3831748

C:Superfamily: low density lipoprotein receptor-related protein; EGF homology; LDL recept

F:605-640/Domain: EGF homology <EGF>

F:1259-1295/Domain: LDL receptor ligand-binding repeat homology <LDL1>

F:1298-1332/Domain: LDL receptor ligand-binding repeat homology <LDL2>

F:1336-1370/Domain: LDL receptor ligand-binding repeat homology <LDL3>

Query Match 100.0%; Score 8736; DB 2; Length 1615;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1614; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	1	MEAAPGPPWPLLLLLLLLLLALCGCPAPAAASPLLLFANRRDRLVLDAGGVKLESTIVVS	60
Db	1	MEAAPGPPWPLLLLLLLLLLALCGCPAPAAASPLLLFANRRDRLVLDAGGVKLESTIVVS	60
Qy	61	GLEDAAAVDFQFSKGVYVTDVSEEAIKQTYLNQTGAQVQNVVIGSLVSPGLACDWGK	120
Db	61	GLEDAAAVDFQFSKGVYVTDVSEEAIKQTYLNQTGAQVQNVVIGSLVSPGLACDWGK	120
Qy	121	KLYWTDSETNRIEVANLNGTSRKVLFWQDLDPRAIALDPAGHYWYTDWGETPRIERAG	180
Db	121	KLYWTDSETNRIEVANLNGTSRKVLFWQDLDPRAIALDPAGHYWYTDWGETPRIERAG	180
Qy	181	MDGSTRKLIIVDSDIYWPNGLTIDLEEQKLYWADAKLSFIHRANLDGSPQKRVGSLTTP	240
Db	181	MDGSTRKLIIVDSDIYWPNGLTIDLEEQKLYWADAKLSFIHRANLDGSPQKRVGSLTTP	240
Qy	241	FALTSLSGDTLYTWDQWTRSHACNKRKTGGRKEILSALYSPMDIQVLQVLSQERQPPFHTRC	300
Db	241	FALTSLSGDTLYTWDQWTRSHACNKRKTGGRKEILSALYSPMDIQVLQVLSQERQPPFHTRC	300

Db 241 FALTLSGDTLYWTDNQTRSIHACNKRKTKGKRKEIILSALYSMPMDIQVLSQERQPPFHTTCE 300
Qy 301 EDNGGCSHLCLLSPSEPFYTCACPTGVQLQDNNGRTCKAGAEVLLARTRDLRLISLDT 360
Db 301 EDNGGCSHLCLLSPSEPFYTCACPTGVQLQDNNGRTCKAGAEVLLARTRDLRLISLDT 360
Qy 361 DFTDVLQVDDIRHAIADYDPLEGYVYVWTDDEVRAIRRAYLDGSGAQTLVNTEINDPDG 420
Db 361 DFTDVLQVDDIRHAIADYDPLEGYVYVWTDDEVRAIRRAYLDGSGAQTLVNTEINDPDG 420
Qy 421 IAVDMVARNLYWTDGTDRIEVTRNGTSRKLTVSEDLDEPRAIALHPVGMGLMYWTDWGE 480
Db 421 IAVDMVARNLYWTDGTDRIEVTRNGTSRKLTVSEDLDEPRAIALHPVGMGLMYWTDWGE 480
Qy 481 NPKIECANLQGERRVLVNASLGNWGLALDQEGKLYWDAKTDKIEVINVDGTKRRL 540
Db 481 NPKIECANLQGERRVLVNASLGNWGLALDQEGKLYWDAKTDKIEVINVDGTKRRL 540
Qy 541 LEDKLPHIFGFTLLGDFYITWQRRSIEVHKVKASRDVLIIDQLPDLMLKAVNAVW 600
Db 541 LEDKLPHIFGFTLLGDFYITWQRRSIEVHKVKASRDVLIIDQLPDLMLKAVNAVW 600
Qy 601 GTNPCADRNGGCSHLCTFTPHATRCGCPGLELLSDMTKICVPRAFVFTSRAAHRISL 660
Db 601 GTNPCADRNGGCSHLCTFTPHATRCGCPGLELLSDMTKICVPRAFVFTSRAAHRISL 660
Qy 661 ETNNNDVAIPITGVKEASALDPDVSNHHIYWDVSLKTSIRAFMNGSSVEHVFEGLDYP 720
Db 661 ETNNNDVAIPITGVKEASALDPDVSNHHIYWDVSLKTSIRAFMNGSSVEHVFEGLDYP 720
Qy 721 EGMADVMMKGLYWDATGTNRIEVARLDGQROVLVWDLNPRSLALDPTKGYIYTW 780
Db 721 EGMADVMMKGLYWDATGTNRIEVARLDGQROVLVWDLNPRSLALDPTKGYIYTW 780
Qy 781 GKGPRIVAFNDGTNCMTLVKVGANDLTIDYADQRLYTDLDTNMIENSNMLGOERVV 840
Db 781 GKGPRIVAFNDGTNCMTLVKVGANDLTIDYADQRLYTDLDTNMIENSNMLGOERVV 840
Qy 841 IADDLPHFPGLTQYSDIYIYWDNHLNHSIERADKTSGRNRTLQGHLDPMVDILVFHSRQ 900
Db 841 IADDLPHFPGLTQYSDIYIYWDNHLNHSIERADKTSGRNRTLQGHLDPMVDILVFHSRQ 900
Qy 901 DGLNDCHNNGCCGOLCLAIPEGHRCGCASHYTLDPSSRNCSPPTTLLFSQKSAISMI 960
Db 901 DGLNDCHNNGCCGOLCLAIPEGHRCGCASHYTLDPSSRNCSPPTTLLFSQKSAISMI 960
Qy 961 PDDQHSPLILPLHGLRNVAKIDYDPLDKFYIYWDGRQNIKRAKDDGTQPFVLTSLSGQ 1020
Db 961 PDDQHSPLILPLHGLRNVAKIDYDPLDKFYIYWDGRQNIKRAKDDGTQPFVLTSLSGQ 1020
Qy 1021 NPDQPHDLSIDYISRTLFWTCEATNTINVRHLSGEAMGVVLRGDRDKPRAIVVNAERG 1080
Db 1021 NPDQPHDLSIDYISRTLFWTCEATNTINVRHLSGEAMGVVLRGDRDKPRAIVVNAERG 1080
Qy 1081 LYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLFWVDADLRKIESCD 1140
Db 1081 LYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLFWVDADLRKIESCD 1140
Qy 1141 LSGANRLTLEDANIVQPLGLTILGKHLWIDRQOQMIEVEKTTGDKRTRIQGRVAHLTG 1200
Db 1141 LSGANRLTLEDANIVQPLGLTILGKHLWIDRQOQMIEVEKTTGDKRTRIQGRVAHLTG 1200
Qy 1201 IHAVEEVSLEEFSAHPCARDNGGCSHICIAKGDGTPRCSCPVLVQLNLLTCGEPTCS 1260
Db 1201 IHAVEEVSLEEFSAHPCARDNGGCSHICIAKGDGTPRCSCPVLVQLNLLTCGEPTCS 1260
Qy 1261 PDQFACATGEIDCIPGAWRCDFPECDQSDQSEEGCPVCSAAQFPCCARQCQVDLRLCDDGE 1320
Db 1261 PDQFACATGEIDCIPGAWRCDFPECDQSDQSEEGCPVCSAAQFPCCARQCQVDLRLCDDGE 1320
Qy 1321 ADCQDRSEADCAICLPNQFRCSAGQCCLIKQCCDSDPDICDSDGSDLMCEBITKPPSDDS 1380
Db 1321 ADCQDRSEADCAICLPNQFRCSAGQCCLIKQCCDSDPDICDSDGSDLMCEBITKPPSDDS 1380

Qy 1381 PAHSSAIGPVGIIILSLFVMGVYFVQORVVCORYAGANGPFPHEYYSGTGHVPLNFIAP 1440
Db 1381 PAHSSAIGPVGIIILSLFVMGVYFVQORVVCORYAGANGPFPHEYYSGTGHVPLNFIAP 1440
Qy 1441 GGSQHGFPTGIACGKSMSSVSLMGGRGVPLYDRNHVGTGASSSSSSSTKATLPPILNP 1500
Db 1441 GGSQHGFPTGIACGKSMSSVSLMGGRGVPLYDRNHVGTGASSSSSSSTKATLPPILNP 1500
Qy 1501 PPSPATPPLYNMDFYSSNIPATAPRVPYIIRGMAPPPTTTCSTDVCDSDYSASRWKAS 1560
Db 1501 PPSPATPPLYNMDFYSSNIPATAPRVPYIIRGMAPPPTTTCSTDVCDSDYSASRWKAS 1560
Qy 1561 KYILDLNSDSDYPPTPHPSQYLSAEDSCPPSPATERSYFHLFPPTPPTDSS 1615
Db 1561 KYILDLNSDSDYPPTPHPSQYLSAEDSCPPSPATERSYFHLFPPTPPTDSS 1615

RESULT 2
JE0272
low density lipoprotein receptor-related protein 6 - human
C:Species: Homo sapiens (man)
C:Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 09-Jul-2004
C:Accession: JE0272
R:Brown, S.D.; Twells, R.C.J.; Hey, P.J.; Cox, R.D.; Levy, E.R.; Soderman, A.R.; Metzker, B. Biochem. Biophys. Res. Commun. 248, 879-888, 1998
A:Title: Isolation and characterization of LRP6, a novel member of the low density lipoprotein receptor family
A:Reference number: JE0272; MUID:98369644; PMID:9704021
A:Accession: JE0272
A:Molecule type: mRNA
A:Residues: 1-1613 <BRO>
A:Cross-references: UNIPROT:O75581; GB:AF074264; NID:G3462526; PIDN:AAC33006.1; PID:G3462526
C:Genetics:
A:Gene: LRP6
A:Map position: 12p11-12p13
C:Superfamily: low density lipoprotein receptor-related protein; EGF homology; LDL receptor
F:286-323/Domain: EGF homology <EGF1>
F:592-627/Domain: EGF homology <EGF2>
F:1249-1285/Domain: LDL receptor ligand-binding repeat homology <LDL1>
F:1288-1322/Domain: LDL receptor ligand-binding repeat homology <LDL2>
F:1326-1360/Domain: LDL receptor ligand-binding repeat homology <LDL3>

Query Match 70.4%; Score 6153.5; DB 2; Length 1613;
Best Local Similarity 70.2%; Pred. No. 0;
Matches 1134; Conservative 205; Mismatches 256; Indels 21; Gaps 12;

Qy 15 LLLLLALLCGCPAPAAAPLIPANRRDVRIVDAGGVKLESTIVVSGLEDAADVDFFSK 74
Db 4 VLRSLLACSCFVLLRAA-PLLLVYANRRDLRLVDTNKENATIVVSGLEDAADVDFFSH 62
Qy 75 GAVYTWTDVSEAIKQTYLNTQGAQVQVNVISGLVSPDGLACDWVVKGLYWTDSFTRIEV 134
Db 63 GLIYMSDVSEAIKRTFNKT-ESVQVNVVSGLLSPDGLACDWLGEKLYWTDSFTRIEV 121
Qy 135 ANLNGTSRKLFWQDLDPRAIALDPAHGYWYWTWGTETPRIERAGDGGSTRKLIIVSDI 194
Db 122 SNLDGSLRKLFWQELDPRAIALDPSGPFMYWTWGEVFKIERAGDGGSTRFIINSEI 181
Qy 195 WYPNGLTIDLEEOKLYWADAKLSFIHRANDLGGSPROKRVGSLTHFPALTLSGDTLYWTD 254
Db 182 WYPNGLTIDYEEQKLYWADAKLFIHNSLDNTGRVAVKSLPHFPALTILFEDILYWTD 241
Qy 255 WQTRSIHACNKRKTKGKRKEIILSALYSMPMDIQVLSQERQPPFHTTCEEDNGCCHLCLSP 314
Db 242 WSTHISILACNKYTGEGRLREIHSIDFSPMDIHAFSQQPNATNPGCIDNGCCHLCLMSP 301
Qy 315 SEPYTTCACGTGVLQDNGRTCKAGAEVLLARTRDLRLISLDTDFTDVLQVDDIRH 374
Db 302 VKPPYQACPTGVKLENGKTKDGAATELLARTRDLRLISLDTDFTDVLQVDDIRH 361
Qy 375 AIAIDYDPLEGYVYVWTDDEVRAIRRAYLDGSGAQTLVNTEINDPDGIAVDMVARNLYWTD 434

Db 362 AAIADYDPVEGIYIYWTDDVEAIRRSFIDGSGSQFVVTQAIAHPDGIADVAVARNLYWTD 421
Qy 435 TGTDRIEVTRNGTSRKILVSEDLDEPRAIALHPVMGLMYWTDWGENPKIECANLDGOER 494
Db 422 TGTDRIEVTRNGTWKILISEDLDEPRAIVLDPWGVYMYWTDWGEIPKIERAALDGSDR 481
Qy 495 RVLVNASLWPNGLALDQEGKLYWGDAKTDKIEVINVDGTRKRTLLEDKLPHPFGFTLL 554
Db 482 VVLVNTSLGWPNGLALDYDEGKLYWGDAKTDKIEVMNTDGTGRRLVVEDKIPHPFGFTLL 541
Qy 555 GDFLYWTDWQBSRTERVHKVYKASRDVVIDQLPDLMLKAVNAVGVNTPCADRNGGCSH 614
Db 542 GDIYVWTDWQBSRTERVHKVYKASRDVVIDQLPDLMLKATNVHVRVIGSNPCAEENGCSH 601
Qy 615 LCFPTPHATRCGCPGLELLSDMTKICIVPEAFVFTSRAAIHRISLETNNNDVAIPITGV 674
Db 602 LCLYRPOGLRCACPIGELISDMTKICIVPEAFLEFSRADIRISLETNNNDVAIPITGV 661
Qy 675 KEASALDFDVSNHHIYWTDVSLKTIISRAFPMNGSSVHVVEFGLDYPEGMAVDMWGNKLYW 734
Db 662 KEASALDFDVTDNRNIYWTDISLKTISRAFPMNGSALEHVVEFGLDYPEGMAVDMWGNKLYW 721
Qy 735 ADTGTRIEVARLDQFQVVLWRDLNPRSIALDPTKGYIYWTWEGKPKRIVAFMDGT 794
Db 722 ADTGTRIEVSKLDGQHRQVVLWKDLSPRALALDPAEGFMYTWEWGKPKRIDRAANDGS 781
Qy 795 NCMTLVKVGANDLTIDYADORLYWTDLTNMTIESSNMLGOERVVIADDLPHRPGLTQY 854
Db 782 ERTTLVNVVGRANGLTIDYAKRRLYWTDLTNLTIESSNMLGNEVVIADDLPHRPGLTQY 841
Qy 855 SDIYVWTDWNLHISIERADKTSGRNRTIIOGHLDFVMDILVPHSSRQGLDCHMNNQCG 914
Db 842 QDIYVWTDWMSRISIERANKTSQNRITIIQGLHYVMDILVPHSSRQGWNECASNHGCS 901
Qy 915 QLCIAIP-GGHRGCGASHYTLDPSSNCSPTTFLFSQKSAISRMIPDDQHSPLILPL 973
Db 902 HLCIAVFGVGVCGPAHYSLNADNRCTASPTTFLFSQKSAINRMVDEQSSDIIILPI 961
Qy 974 HGLRNVRAIDYDPLDKFYWVDGRON-IKRAKDGTQFP-VLTSLSQONPDPRPHDLSI 1031
Db 962 HSLRNVRAIDYDPLDKQYIWDQRNIMIRKAQEDGSGQFTVVVSSVPSQNLIEIOPYDLSI 1021
Qy 1032 DIYSRTLFTWCEATNTNVHRLSGEAMGVVLGRDRKPRAIVVNAERGILYFTNNQDRAA 1091
Db 1022 DIYSRYIYWTCEATNVINVTELDGRSVGLVKGQDRPRAIVVNPKEGYMYFTNLQERSP 1081
Qy 1092 KIERAALDGTREVLFTGLIRPVALVVDNTLGLKLFWVDADLKRIESCDSGANRLTLED 1151
Db 1082 KIERAALDGTREVLFFSGLSKPIALALDSRLGLFWADSDLRRIESSDLSGANRIVLED 1141
Qy 1152 ANIVQPLGLTLGKHLWIDRQOQMIERVEKTTGDKRTRIOGRVAHLTGTHAVEVSLEE 1211
Db 1142 SNILQPVGLTVFENWLYWIDRQOQMIKIDMTGREGRTKVQARQAQLSDIHAVKELNLQE 1201
Qy 1212 FSAHPCARDNGCGSHICTAGDGTGPRSCPVHLVLLQNLITCGEPPTCSPDQFACATGEI 1271
Db 1202 YRQHPCAQDNGCGSHICLVKGDGTTRCSPHVLVLLQDELSCGEPPTCSPOQFTCTGGEI 1261
Qy 1272 DCIPCAWRCDGFPBCDDQSDDEEGCPVCSAAQFPCCARGQCVDLRLRCDEADCDQSDRDEAD 1331
Db 1262 DCIPVAMRCDFTECEDHSDDELNCPVCSQSFQCCASQCIDGALRCNGDANCQKSDKEN 1321
Qy 1332 CDALCLNQCPCASQCVLIIKQOQSDPDCIDGSDLMCEITKPPSDSDSPAHSAGPVI 1391
Db 1322 CEVLICLDIOPRCANGQCIGRHKKCDHNVDCKDELDCYPTTEP---APQATNTVGSVI 1378
Qy 1392 GIILSLFWMGVYVFCORVVCORVAGANGPPFHEY-VSGTPHVPLNFTAPGSGHQGPFTG 1450
Db 1379 GVIITVIFSGTVYVICRMLCFPMKGDGETWNTVYVHGPAVPLGVYVPHPSLSGSLPG 1438
Qy 1451 IACGKSMSSVSLMGGRGVPLYDRNHVTVGASSSSSSTKATLYPPIINPPSPSPATPDSL 1510
Db 1439 MSRGKSMISSLSIMGSSGPP-YDRAHVTGASSSSSSSTKGTYPPIALINPPSPSPATPDSH 1497

Qy 1511 YNMDMFYSSNIPATAR--PYRPIIRGMAPPPTPCSTDVDCSDYSASR-----WKASKYY 1563
Db 1498 YTFEFGYSSNSPSTHRSYSYRPIYRHFAPPTTFCSTDVDCSDYAPSRMTSVATARGYT 1557
Qy 1564 LDLSNSDSDPPPPPTPHSQVLSAB---DSCPPSPATERSY-FHLPPPPSPCTOSS 1615
Db 1558 SDLNYDSEPVPPTTPRSQVLSABENYESCPPSPYTERSYSHHLYPPPPSPCTOSS 1613

RESULT 3

JE0273
low density lipoprotein receptor-related protein 6 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 09-Jul-2004
C:Accession: JE0273
R:Brown, S.D.; Twells, R.C.J.; Hey, P.J.; Cox, R.D.; Levy, E.R.; Soderman, A.R.; Metzker, Biochem. Biophys. Res. Commun. 248, 879-888, 1998
A:Title: Isolation and characterization of LRP6, a novel member of the low density lipoprotein receptor family
A:Reference number: JE0272; MUID:98369644; PMID:9704021
A:Accession: JE0273
A:Molecule type: mRNA
A:Residues: 1-1613 <BRO>
A:Cross-references: UNIPROT:O88572; GB:AF074265; NID:G3462528; PIDN:AAC33007.1; PID:G3462528
C:Genetics:
A:Gene: Lrp6
A:Map position: 6
C:Superfamily: low density lipoprotein receptor-related protein; EGF homology; LDL receptor
F:286-323/Domain: EGF homology <EGF1>
F:592-627/Domain: EGF homology <EGF2>
F:1207-1243/Domain: EGF homology <EGF3>
F:1249-1285/Domain: LDL receptor ligand-binding repeat homology <LDL1>
F:1288-1322/Domain: LDL receptor ligand-binding repeat homology <LDL2>
F:1326-1360/Domain: LDL receptor ligand-binding repeat homology <LDL3>

Query Match 69.8%; Score 6097.5; DB 2; Length 1613;
Best Local Similarity 69.3%; Pred. No. 0;
Matches 1120; Conservative 216; Mismatches 259; Indels 21; Gaps 12;

Qy 15 LLLLLLLALCGCPAPAAASPLLFANRRDRVLVDAGGVKLESTIVVSGLEDAADVDFQFSK 74
Db 4 VLSRLTACSECVLLRAA-PLLLYANRRDLRLVDATNGKENATIVVGGLEDAAADVDFEGH 62
Qy 75 GAVVTVDSSEAIKQTYLNTGAAVQNVWISGLVSPDGLACDWMVGKLYWTDSETRLEV 134
Db 63 GLIYWSVSEAIKRTBFNKS-ESVQNVVSGLLSPGLACDNLGKLYWTDSETRLEV 121
Qy 135 ANLNGTSRKVLFWQDLDPRAIALDPAHGYMYWTDWGETPRIERAGMDGSTRKIVSDI 194
Db 122 SNLDGSLRKVLFWQDLDPRAIALDPSGFMWYTDWGEVPKIERAGMDGSSRFVIINTEI 181
Qy 195 YWPNGLTIDLEEQKLYWADAKLSFIHRANLDGSPRKVVEGSLTHPPALITLSGDTLYWTD 254
Db 182 YWPNGLTIDYQERKLYWADAKLFIHNSLDGNTNRQAVVKGSLPFPFALTIFEDTLYWTD 241
Qy 255 WQTESIHACNKRKTKGKKEILSALYSPMDIQVLSQERQPPFHTTRCEDNGCASHLLCLSP 314
Db 242 WNTSHIACNKYTGEGUREIHSNIFSPMDIHAFQQQPNATNFCGIDNGCASHLLCLMSP 301
Qy 315 SEPYTTCACPTGVQDQNGRTCKAGAEVLLAARTDLRRLSLDPTDFDTIVLQVDDIRH 374
Db 302 VKPFYCACPTGVKLMENKTKDGATELLLLAARTDLRRLSLDPTDFDTIVLQVDDIRH 361
Qy 375 AIAIDYDPLGEGYVWTDDEVRAPRAVLDGSGAGTIVNTEINDPDGTAADVARNLYWTD 434
Db 362 AIAIDYDVEGYIYWTDEVRAIRRSFIDGSGSQFVVTQAIAHPDGIADVAVARNLYWTD 421
Qy 435 TGTDRIEVTRNGTSRKILVSEDLDEPRAIALHPVMGLMYWTDWGENPKIECANLDGOER 494
Db 422 TGTDRIEVTRNGTWKILISEDLDEPRAIVLDPWGVYMYWTDWGEIPKIERAALDGSDR 481
Qy 495 RVLVNASLWPNGLALDQEGKLYWGDAKTDKIEVINVDGTRKRTLLEDKLPHPFGFTLL 554

Db 482 VVLVNTSLGWPNGLALDYDEGTIYWGDAKTDKIEVMNTDGTGRRVLVLEKIPHIFGFTLL 541
QY 555 GDFIYWTWQRESIERVHKVKSARDVIIDQLPDLMLGLKAVNVKVGTVNCPADRNGGSGH 614
Db 542 GDYVYWTWQRESIERVHKRSAREVIIIDQLPDLMLGLKATSVHVRVIGSNPCAEADNGGCSH 601
QY 615 LCFPTPHATRCGCPGLELLSMDKTCIPEAEALVFTSRAAIIHRISETNNNDVAIPLTVG 674
Db 602 LCLYRPOGLRCACPIGFIIDGMKTCIPEAEALLFSRRADIRRISETNNNDVAIPLTVG 661
QY 675 KEASALDFDVSNNHIYWTVDLSUKTISRAPMGSSVEHVFEGLDYPEGMVMDMMGNLYW 734
Db 662 KEASALDFDVTIDNRYYWTVDLSUKTISRAPMGSSALEHVFEGLDYPEGMVMDLGNLYW 721
QY 735 ADTGNNRLEVARLDGQFVLYWRDLNPRSLALDPTKGYIYWTWGGKPRIVRAFMDGT 794
Db 722 ADTGNNRLEVSCLDGQHRQLVWKDLSPRALDPAEGFMYWTWGGKPKIDRAMDGS 781
QY 795 NCWTLVVKVGRANDLTIDYADQRLYWTDLTDNMIESNNMLGOERVVVIADDLPHPPGLTQY 854
Db 782 BRTTLVNVGRANGLTIDYAKERLYWTDLTDNLIESDMLGNREVIADDLPHPPGLTQY 841
QY 855 SYIYVTDNLHLSIERAKTSGRNTLIQGHLDYFVMDILVFHSSRODGLNDCMHNQCG 914
Db 842 QDIYIYVTDWSRRSIRERANKTSQNRITIIQGHLDYVMDILVFHSSRQAGWNECASSNGHCS 901
QY 915 QLCLAIIP-GGHCRCASHVTLDPSSRNCSPPTFLIFSQKSAISRMI PDDQHSPLLIPL 973
Db 902 HLCLAVPVGFGVCGCPAHYSLNADNRCTSAPEFLIFSQKSAINRMVDEOQSPDIILPI 961
QY 974 HGLRNVKAIDYDPLDKFYIYVNDGRQN-IKRAKMDGTQPF-VLTLSQCGNPDROPHDLSI 1031
Db 962 HSLRNVRAIDYDPLDKQLYWIDSRQNSIRKAHEDGGQGFNVVANSVANQNLIEIYDLSI 1021
QY 1032 DIYSRTLFWTCBATNTINVHRLSGBAMGVVLGRDRKPRAIVVNAERGYLYFTNNQDRAA 1091
Db 1022 DIYSRYIYWTCEATNVIDTRLDRSGVGLKGEQDRPRAIVVNEKGYMYFTNLQERSP 1081
QY 1092 KIERAALDGTREVLFTTGLIRPVALVVDNTLGLFWVDADLKRIESCDSLGNRLTLED 1151
Db 1082 KIERAALDGTREVLFTTGLIRPVALVVDNTLGLFWVDADLKRIESCDSLGNRLTLED 1141
QY 1152 ANIVQPLGLTILGKHLIYWDROQQMIERVEKTTGDKRTRIQGRVAHLTGIIHAEVSELS 1211
Db 1142 SNILQPVGLTFENWLYWIDROQQMIEXIDMTGREGRTKVQARIAQLSDIIHAEVSELS 1201
QY 1212 FSAHPCARDNGGCSHICTAKGDTGPRCSPVHLVLLQNLITCGEPPTCSPQFACATGEI 1271
Db 1202 YROHPCARDNGGCSHICLVKGDGTTTCSPVHLVLLQDELSCGEPPTCSPQFCTGTGDI 1261
QY 1272 DCIPGAWRCDFPECCDOSDEEGCPVCSAAQPPCARGOCVDLRLRCDEADCODESDEAD 1331
Db 1262 DCIPVAMRCDFTECEDHSDLCNCPVCSQESQFQASGQCIDGALRCNGDANCQDKSDKN 1321
QY 1332 CDAICLPNQCASCQCQVLIKQCCDFPDCIDGSDLMCEITKPPSPATDPSLYNMDFYSSNIATARP 1391
Db 1322 CEVLCLIDQFRANCQCQVGHKKCDHSDVCSDRSDELDCYTEBP--APQATVTVGSVI 1378
QY 1392 GILLSLFVWGVIYVFCQVRVQVRVAGANGPPHPPHVEVSGTP-HVPLNFTAPGSGHQPTG 1450
Db 1379 GVIVTIFVSGTIYFTICRMCLCPMKMGDGETTNDYVHVSASVPLGYVPHPSLSGSLPG 1438
QY 1451 IACGKSMSSVSLMGGRGVPLYDRNHNVTGASSSSSTKATLYPPIINPPSPATDPSL 1510
Db 1439 MSRGKSMISSLSIMGGSGGP- YDRAHVTGASSSSSTKATLYPPIINPPSPATERSH 1497
QY 1511 YNMDMFYSSNIATAR--PYRPIYIRGMAPPTPCSTDVCDSDYSASR-----WKASKY 1563
Db 1498 YTMREGYSSNSPSTHRSYSYRPSYRHPAPTPTPCSTDVCDSDYAPSRMTSVATAGYT 1557
QY 1564 LDLNSDSDPYPPPTPHSQYLSAE---DSCPPSPATERSY-FHLFPPPPSPCTDSS 1615
Db 1558 SDVNYDSEFVPPPTPRSQYLSAENYENSCPPSPYTERSYGHLYPPPPSPCTDSS 1613

RESULT 4
JB0373

low density lipoprotein receptor related protein - rabbit

C:Species: Oryctolagus cuniculus (domestic rabbit)

C:Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 09-Jul-2004

C:Accession: JB0373

R:Kim, D.; Inagaki, Y.; Suzuki, T.; Ioka, R.X.; Yoshioka, S.Z.; Magoori, K.; Kang, M.; Ch

J. Biochem. 124, 1072-1076, 1998

A:Title: A new low density lipoprotein receptor related protein, LRP5, is expressed in he

A:Reference number: JB0372; MUID:99054722; PMID:9832610

A:Accession: JB0373

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-527 <KIM>

A:Cross-references: UNIPROT:O77501; DBJ:AB017499

C:Superfamily: low density lipoprotein receptor-related protein; EGF homology; LDL recept

F:211-247/Domain: LDL receptor ligand-binding repeat homology <LDL1>

F:250-284/Domain: LDL receptor ligand-binding repeat homology <LDL2>

F:288-322/Domain: LDL receptor ligand-binding repeat homology <LDL3>

Query Match

Best Local Similarity 30.4%; Score 2654; DB 2; Length 527;

Matches 482; Conservative 19; Mismatches 25; Indels 0; Gaps 0;

QY 1049 NVHRLSGHMGVVLGRDRKPRAIVVNAERGYLYFTNNQDRRAAKIERAALDGTREVLFT 1108

Db 1 NVQRLSGEAVGVVLGRDRKPRAIVVNAERGHLYFTNNQDRRAAKIERAALDGTREVLFT 60

QY 1109 TGLIRPVALVVDNTLGLFWVDADLKRIESCDSLGNRLTLEDANIYQPLGTLILGKHL 1168

Db 61 TGLIRPVALVVDNAGLKLFWVDADLKRIESCDSLGNRLTLEDASIVQPVGLVILGRHL 120

QY 1169 WIDRQQMIERVEKTTGDKRTRIQGRVAHLTGIIHAEVSELSLEEFSAHPCARDNGGCSHIC 1228

Db 121 WADRQQMIERVEKATGDKRTVQGRVAHLTGIIHAEVSELSLEEFSAHPCARDNGGCSHIC 180

QY 1229 IAKGDTGPRCSPVHLVLLQNLITCGEPPTCSPQFACATGEIDICIPGAWRCDFPECCD 1288

Db 181 IAKGDTGPRCSPVHLVLLQNLITCAEPPTCSPQFACATGEIDICIPGAWRCDFPECCD 240

QY 1289 QSDREGCPVCSAAQPPCARGOCVDLRLRCDEADCODESDEADCAICLPNQCASCQC 1348

Db 241 QSDREGCPVCSAAQPPCARGOCVDLRLRCDEADCHDHSDEADCAICLPNQCASCQC 300

QY 1349 VLIKQCCDFPDCIDGSDLMCEITKPPSPATDPSLYNMDFYSSNIATARP 1408

Db 301 LLIRQCCDFPDCVGDSDLMCEITRPPVDDIPAHSSAIGPVGIIILSLFVWGVIYVFCQ 360

QY 1409 RVVQVRVAGANGPPHPPHVEVSGTPHVPINFTAPGSGHQPTGACGKSMSSVSLMGGR 1458

Db 361 RVVQVRVAGANGPPHPPHVEVSGTPHVPINFTAPGSGHQPTGACGKSMSSVSLMGGR 420

QY 1469 GVPLYDRNHNVTGASSSSSTKATLYPPIINPPSPATDPSLYNMDFYSSNIATARP 1528

Db 421 AVPLIERNHNVTGASSSSSTKASLYPQILNPPSPATDPSLYNMDFYIPANIPTATPY 480

QY 1529 RPYIIRGMAPPTPCSTDVCDSDYSASRWKASKYLLDNLSDSDPY 1574

Db 481 RSYVIRGMAPPTPCSTDVCDSDYSASRWKASKYLLDNLSDSDPY 526

RESULT 5
S02392

alpha-2-macroglobulin receptor precursor - human

N:Alternate names: CD91; LDL receptor-related protein 1; low density lipoprotein receptor

C:Species: Homo sapiens (man)

C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 09-Jul-2004

C:Accession: S02392; S30207; I37998; A39210; S42538

R:Herz, J.; Hamann, U.; Rognan, S.; Myklebost, O.; Gausepohl, H.; Stanley, K.K.

EMBO J. 7, 4119-4127, 1988

A:Title: Surface location and high affinity for calcium of a 500-kd liver membrane protei

A:Reference number: S02392; MUID:89210795; PMID:3266596
 A:Accession: S02392
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-4544 <HR>
 A:Cross-references: UNIPROT:Q07954; EMBL:X13916; NID:G34338; PIDN:CAA32112.1; PID:G34339
 R:Kristensen, T.
 submitted to the EMBL Data Library, October 1990
 A:Reference number: S30027
 A:Accession: S30027
 A:Molecule type: mRNA
 A:Residues: 3275-3864 <KRI>
 A:Cross-references: EMBL:X55077
 R:Herz, J.; Kowal, R.C.; Goldstein, J.L.; Brown, M.S.
 EMBO J. 9, 1769-1776, 1990
 A:Title: Proteolytic processing of the 600 kd low density lipoprotein receptor-related protein
 A:Reference number: S12538; MUID:90269210; PMID:2112085
 A:Contents: annotation; site of proteolytic cleavage
 R:Kutt, H.; Herz, J.; Stanley, K.K.
 Biochim. Biophys. Acta 1009, 229-236, 1989
 A:Title: Structure of the low-density lipoprotein receptor-related protein (LRP) promote
 A:Reference number: I37998; MUID:90089395; PMID:2597675
 A:Accession: I37998
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-11 <RES>
 A:Cross-references: EMBL:X15424; NID:G34408; PIDN:CAA33464.1; PID:G34409
 R:Strickland, D.K.; Ashcom, J.D.; Williams, S.; Burgess, W.H.; Migliorini, M.; Argraves, J. Biol. Chem. 265, 17401-17404, 1990
 A:Title: Sequence identity between the alpha2-macroglobulin receptor and low density lipoprotein receptor
 A:Reference number: A39210; MUID:91009181; PMID:1698775
 A:Accession: A39210
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 150-166;234-238, 'X', 240-245, 'X', 247-252; 'G', 686-695;902-916;1096-1109; 'S', 1711-1712
 C:Genetics:
 A:Gene: GDB:LRP1; APR; LRP; A2MR
 A:Cross-references: GDB:119694; OMIM:107770
 A:Map position: 12q13.1-12q13.3
 C:Complex: The alpha-2-macroglobulin receptor complex consists of noncovalently-associated protein (see PIR:A39875).
 C:Superfamily: alpha-2-macroglobulin receptor; EGF homology; LDL receptor ligand-binding
 C:Keywords: beta-hydroxyasparagine; beta-hydroxyaspartic acid; calcium binding; glycoprotein
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-3943/Product: alpha-2-macroglobulin receptor 515K chain #status predicted <515K>
 F:27-64/Domain: LDL receptor ligand-binding repeat homology <LDL1>
 F:72-108/Domain: LDL receptor ligand-binding repeat homology <LDL2>
 F:115-148/Domain: EGF homology <EG1>
 F:154-188/Domain: EGF homology <EG2>
 F:198-239/Domain: LDL receptor WYTD-containing repeat homology <YW01>
 F:240-281/Domain: LDL receptor WYTD-containing repeat homology <YW02>
 F:292-334/Domain: LDL receptor WYTD-containing repeat homology <YW03>
 F:335-378/Domain: LDL receptor WYTD-containing repeat homology <YW04>
 F:379-420/Domain: LDL receptor WYTD-containing repeat homology <YW05>
 F:421-468/Domain: LDL receptor WYTD-containing repeat homology <YW06>
 F:478-519/Domain: EGF homology <EG3>
 F:571-613/Domain: LDL receptor WYTD-containing repeat homology <YW07>
 F:614-659/Domain: LDL receptor WYTD-containing repeat homology <YW08>
 F:660-710/Domain: LDL receptor WYTD-containing repeat homology <YW09>
 F:711-752/Domain: LDL receptor WYTD-containing repeat homology <YW10>
 F:753-799/Domain: LDL receptor WYTD-containing repeat homology <YW11>
 F:807-842/Domain: EGF homology <EG4>
 F:854-890/Domain: LDL receptor ligand-binding repeat homology <LDL3>
 F:893-931/Domain: LDL receptor ligand-binding repeat homology <LDL4>
 F:936-971/Domain: LDL receptor ligand-binding repeat homology <LDL5>
 F:976-1011/Domain: LDL receptor ligand-binding repeat homology <LDL6>
 F:1015-1051/Domain: LDL receptor ligand-binding repeat homology <LDL7>
 F:1062-1097/Domain: LDL receptor ligand-binding repeat homology <LDL8>
 F:1104-1140/Domain: LDL receptor ligand-binding repeat homology <LDL9>
 F:1145-1182/Domain: LDL receptor ligand-binding repeat homology <LDL10>
 F:1185-1221/Domain: EGF homology <EG5>
 F:1227-1261/Domain: EGF homology <EG6>
 F:1269-1308/Domain: LDL receptor WYTD-containing repeat homology <YW12>
 F:1309-1355/Domain: LDL receptor WYTD-containing repeat homology <YW13>
 F:1356-1398/Domain: LDL receptor WYTD-containing repeat homology <YW14>
 F:1399-1445/Domain: LDL receptor WYTD-containing repeat homology <YW15>
 F:1446-1488/Domain: LDL receptor WYTD-containing repeat homology <YW16>
 F:1489-1531/Domain: LDL receptor WYTD-containing repeat homology <YW17>
 F:1540-1578/Domain: EGF homology <EG7>
 F:1583-1626/Domain: LDL receptor WYTD-containing repeat homology <YW18>
 F:1627-1669/Domain: LDL receptor WYTD-containing repeat homology <YW19>
 F:1670-1713/Domain: LDL receptor WYTD-containing repeat homology <YW20>
 F:1714-1753/Domain: LDL receptor WYTD-containing repeat homology <YW21>
 F:1754-1796/Domain: LDL receptor WYTD-containing repeat homology <YW22>
 F:1797-1846/Domain: LDL receptor WYTD-containing repeat homology <YW23>
 F:1850-1886/Domain: EGF homology <EG8>
 F:1934-1976/Domain: LDL receptor WYTD-containing repeat homology <YW24>
 F:1977-2019/Domain: LDL receptor WYTD-containing repeat homology <YW25>
 F:2020-2063/Domain: LDL receptor WYTD-containing repeat homology <YW26>
 F:2064-2105/Domain: LDL receptor WYTD-containing repeat homology <YW27>
 F:2106-2151/Domain: LDL receptor WYTD-containing repeat homology <YW28>
 F:2159-2194/Domain: EGF homology <EG9>
 F:2199-2241/Domain: LDL receptor WYTD-containing repeat homology <YW29>
 F:2253-2294/Domain: LDL receptor WYTD-containing repeat homology <YW30>
 F:2344-2388/Domain: LDL receptor WYTD-containing repeat homology <YW31>
 F:2389-2429/Domain: LDL receptor WYTD-containing repeat homology <YW32>
 F:2430-2473/Domain: LDL receptor WYTD-containing repeat homology <YW33>
 F:2482-2517/Domain: EGF homology <EG10>
 F:2524-2561/Domain: LDL receptor ligand-binding repeat homology <LDL11>
 F:2566-2600/Domain: LDL receptor ligand-binding repeat homology <LDL12>
 F:2605-2639/Domain: LDL receptor ligand-binding repeat homology <LDL13>
 F:2652-2688/Domain: LDL receptor ligand-binding repeat homology <LDL14>
 F:2696-2730/Domain: LDL receptor ligand-binding repeat homology <LDL15>
 F:2734-2769/Domain: LDL receptor ligand-binding repeat homology <LDL16>
 F:2774-2812/Domain: LDL receptor ligand-binding repeat homology <LDL17>
 F:2818-2853/Domain: LDL receptor ligand-binding repeat homology <LDL18>
 F:2858-2897/Domain: LDL receptor ligand-binding repeat homology <LDL19>
 F:2904-2939/Domain: LDL receptor ligand-binding repeat homology <LDL20>
 F:2944-2980/Domain: EGF homology <EG11>
 F:2986-3021/Domain: EGF homology <EG12>
 F:3029-3068/Domain: LDL receptor WYTD-containing repeat homology <YW34>
 F:3069-3113/Domain: LDL receptor WYTD-containing repeat homology <YW35>
 F:3114-3156/Domain: LDL receptor WYTD-containing repeat homology <YW36>
 F:3157-3200/Domain: LDL receptor WYTD-containing repeat homology <YW37>
 F:3201-3241/Domain: LDL receptor WYTD-containing repeat homology <YW38>
 F:3242-3284/Domain: LDL receptor WYTD-containing repeat homology <YW39>
 F:3294-3330/Domain: EGF homology <EG13>
 F:3334-3369/Domain: LDL receptor ligand-binding repeat homology <LDL21>
 F:3374-3408/Domain: LDL receptor ligand-binding repeat homology <LDL22>
 F:3413-3448/Domain: LDL receptor ligand-binding repeat homology <LDL23>
 F:3453-3489/Domain: LDL receptor ligand-binding repeat homology <LDL24>
 F:3494-3531/Domain: LDL receptor ligand-binding repeat homology <LDL25>
 F:3536-3570/Domain: LDL receptor ligand-binding repeat homology <LDL26>
 F:3575-3609/Domain: LDL receptor ligand-binding repeat homology <LDL27>
 F:3613-3647/Domain: LDL receptor ligand-binding repeat homology <LDL28>
 F:3654-3690/Domain: LDL receptor ligand-binding repeat homology <LDL29>
 F:3695-3731/Domain: LDL receptor ligand-binding repeat homology <LDL30>
 F:3741-3776/Domain: LDL receptor ligand-binding repeat homology <LDL31>
 F:3785-3822/Domain: EGF homology <EG14>
 F:3828-3860/Domain: EGF homology <EG15>
 F:3868-3911/Domain: LDL receptor WYTD-containing repeat homology <YW40>
 F:3912-3959/Domain: LDL receptor WYTD-containing repeat homology <YW41>
 F:3944-4420/Product: alpha-2-macroglobulin receptor 85K chain #status predicted <85K>
 F:3944-4420/Domain: 85K chain extracellular #status predicted <EXT>
 F:3970-4012/Domain: LDL receptor WYTD-containing repeat homology <YW42>
 F:4013-4056/Domain: LDL receptor WYTD-containing repeat homology <YW43>
 F:4057-4099/Domain: LDL receptor WYTD-containing repeat homology <YW44>
 F:4100-4142/Domain: LDL receptor WYTD-containing repeat homology <YW45>
 F:4151-4182/Domain: EGF homology <EG16>
 F:4200-4231/Domain: EGF homology <EG17>
 F:4236-4267/Domain: EGF homology <EG18>
 F:4272-4303/Domain: EGF homology <EG19>
 F:4308-4339/Domain: EGF homology <EG20>
 F:4344-4374/Domain: EGF homology <EG21>
 F:4377-4408/Domain: EGF homology <EG22>
 F:4421-4444/Domain: transmembrane #status predicted <TM>

F:662-712/Domain: LDL receptor YWTD-containing repeat homology <YW09>
F:713-754/Domain: LDL receptor YWTD-containing repeat homology <YW10>
F:755-797/Domain: LDL receptor YWTD-containing repeat homology <YW11>
F:805-840/Domain: EGF homology <EG4>
F:852-888/Domain: LDL receptor ligand-binding repeat homology <LDL3>
F:893-929/Domain: LDL receptor ligand-binding repeat homology <LDL4>
F:934-969/Domain: LDL receptor ligand-binding repeat homology <LDL5>
F:974-1009/Domain: LDL receptor ligand-binding repeat homology <LDL6>
F:1013-1049/Domain: LDL receptor ligand-binding repeat homology <LDL7>
F:1060-1095/Domain: LDL receptor ligand-binding repeat homology <LDL8>
F:1102-1138/Domain: LDL receptor ligand-binding repeat homology <LDL9>
F:1143-1180/Domain: LDL receptor ligand-binding repeat homology <LDL10>
F:1183-1219/Domain: EGF homology <EG5>
F:1225-1259/Domain: EGF homology <EG6>
F:1267-1306/Domain: LDL receptor YWTD-containing repeat homology <YW12>
F:1307-1353/Domain: LDL receptor YWTD-containing repeat homology <YW13>
F:1394-1396/Domain: LDL receptor YWTD-containing repeat homology <YW14>
F:1397-1443/Domain: LDL receptor YWTD-containing repeat homology <YW15>
F:1444-1486/Domain: LDL receptor YWTD-containing repeat homology <YW16>
F:1487-1529/Domain: LDL receptor YWTD-containing repeat homology <YW17>
F:1538-1576/Domain: EGF homology <EG7>
F:1581-1624/Domain: LDL receptor YWTD-containing repeat homology <YW18>
F:1625-1667/Domain: LDL receptor YWTD-containing repeat homology <YW19>
F:1668-1711/Domain: LDL receptor YWTD-containing repeat homology <YW20>
F:1712-1751/Domain: LDL receptor YWTD-containing repeat homology <YW21>
F:1752-1794/Domain: LDL receptor YWTD-containing repeat homology <YW22>
F:1795-1842/Domain: LDL receptor YWTD-containing repeat homology <YW23>
F:1846-1882/Domain: EGF homology <EG8>
F:1930-1972/Domain: LDL receptor YWTD-containing repeat homology <YW24>
F:1973-2015/Domain: LDL receptor YWTD-containing repeat homology <YW25>
F:2016-2059/Domain: LDL receptor YWTD-containing repeat homology <YW26>
F:2060-2101/Domain: LDL receptor YWTD-containing repeat homology <YW27>
F:2102-2147/Domain: LDL receptor YWTD-containing repeat homology <YW28>
F:2155-2190/Domain: EGF homology <EG9>
F:2193-2237/Domain: LDL receptor YWTD-containing repeat homology <YW29>
F:2247-2288/Domain: LDL receptor YWTD-containing repeat homology <YW30>
F:2338-2382/Domain: LDL receptor YWTD-containing repeat homology <YW31>
F:2383-2423/Domain: LDL receptor YWTD-containing repeat homology <YW32>
F:2424-2467/Domain: LDL receptor YWTD-containing repeat homology <YW33>
F:2476-2511/Domain: EGF homology <EG10>
F:2518-2554/Domain: LDL receptor ligand-binding repeat homology <LDL5>
F:2560-2594/Domain: LDL receptor ligand-binding repeat homology <LDL6>
F:2599-2633/Domain: LDL receptor ligand-binding repeat homology <LDL7>
F:2646-2682/Domain: LDL receptor ligand-binding repeat homology <LDL8>
F:2690-2724/Domain: LDL receptor ligand-binding repeat homology <LDL9>
F:2732-2767/Domain: LDL receptor ligand-binding repeat homology <LDL10>
F:2772-2810/Domain: LDL receptor ligand-binding repeat homology <LDL11>
F:2816-2851/Domain: LDL receptor ligand-binding repeat homology <LDL12>
F:2856-2895/Domain: LDL receptor ligand-binding repeat homology <LDL13>
F:2902-2936/Domain: LDL receptor ligand-binding repeat homology <LDL14>
F:2941-2977/Domain: EGF homology <EG11>
F:2983-3018/Domain: EGF homology <EG12>
F:3026-3065/Domain: LDL receptor YWTD-containing repeat homology <YW34>
F:3066-3100/Domain: LDL receptor YWTD-containing repeat homology <YW35>
F:3111-3153/Domain: LDL receptor YWTD-containing repeat homology <YW36>
F:3154-3197/Domain: LDL receptor YWTD-containing repeat homology <YW37>
F:3198-3238/Domain: LDL receptor YWTD-containing repeat homology <YW38>
F:3239-3281/Domain: LDL receptor YWTD-containing repeat homology <YW39>
F:3291-3327/Domain: EGF homology <EG13>
F:3331-3366/Domain: LDL receptor ligand-binding repeat homology <LDL15>
F:3371-3405/Domain: LDL receptor ligand-binding repeat homology <LDL16>
F:3410-3445/Domain: LDL receptor ligand-binding repeat homology <LDL17>
F:3450-3486/Domain: LDL receptor ligand-binding repeat homology <LDL18>
F:3491-3528/Domain: LDL receptor ligand-binding repeat homology <LDL19>
F:3533-3567/Domain: LDL receptor ligand-binding repeat homology <LDL20>
F:3572-3606/Domain: LDL receptor ligand-binding repeat homology <LDL21>
F:3610-3644/Domain: LDL receptor ligand-binding repeat homology <LDL22>
F:3651-3687/Domain: LDL receptor ligand-binding repeat homology <LDL23>
F:3692-3728/Domain: LDL receptor ligand-binding repeat homology <LDL24>
F:3738-3774/Domain: LDL receptor ligand-binding repeat homology <LDL25>
F:3783-3820/Domain: EGF homology <EG14>
F:3826-3858/Domain: EGF homology <EG15>
F:3866-3909/Domain: LDL receptor YWTD-containing repeat homology <YW40>

F:3910-3968/Domain: LDL receptor YWTD-containing repeat homology <YW41>
F:3943-4543/Domain: alpha-2-macroglobulin receptor 85k chain #status predicted <85K>
F:3943-4420/Domain: 85k chain extracellular #status predicted <EXT>
F:3969-4011/Domain: LDL receptor YWTD-containing repeat homology <YW42>
F:4012-4055/Domain: LDL receptor YWTD-containing repeat homology <YW43>
F:4056-4098/Domain: LDL receptor YWTD-containing repeat homology <YW44>
F:4099-4141/Domain: LDL receptor YWTD-containing repeat homology <YW45>
F:4150-4181/Domain: EGF homology <EG16>
F:4199-4230/Domain: EGF homology <EG17>
F:4235-4266/Domain: EGF homology <EG18>
F:4271-4302/Domain: EGF homology <EG19>
F:4307-4338/Domain: EGF homology <EG20>
F:4343-4373/Domain: EGF homology <EG21>
F:4376-4408/Domain: EGF homology <EG22>
F:4421-4443/Domain: transmembrane #status predicted <TM>
F:4444-4543/Domain: intracellular #status predicted <INT>
F:116,138,187,276,359,448,731,926,1048,1152,1153,1193,1216,1305,1509,1556,1573,1614,1643,
3485,3659,3786,3837,3952,4074,4124,4178,4278/Binding site: carboxylate (Asn) (covalent)
F:169,2995/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted
F:2955/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status predicted

Query Match 24.9%; Score 2180; DB 1; Length 4543;

Best Local Similarity 34.0%; Pred. No. 2.1e-128; Indels 130; Gaps 49;
Matches 506; Conservative 270; Mismatches 581;

Qy 33 PLLLFANRRDVRVLVDAGGVKLESTIVVSGLEDAADVDFQFSKGVVYTDVSEEAIKQTYL 92
Db 1267 PPIFSNRHEIRDDLH--RGDYSVLVPLENTIALDHLNQSSLYWTDVVDKIYRGKL 1324
Qy 93 NQTGA--AVQNVITSLVSPDGLACVWVKLYWTDSETRIEVANLNGTSRKLFWQDL 150
Db 1325 LENGATSFEVVIQGLATPEGLAVDTIAGNIYVNESNLDQIEVAKLDGWTWRTLLAGDI 1384
Qy 151 DQPAIALLPAHYMYWTDW-GETPRIERAGMDGSTRKI--VDSDIYWPNGLTLDLREQ 207
Db 1385 EHPRAIALLDPYGLTFWTDWDASLPRIEASMSGAGRTTHKTCGSGGWPNGLTVDYLEK 1444
Qy 208 KLYWADAKLSPIHRANLDGSPQKQVWGS--LTHPPFALTLSGDTLYWTDWTRSHACNK 265
Db 1445 RILMDARSDAIYSLVDGTGHIIEVHGHEVLSHPFVTLVGGVEYWTDWTRNTLAKANK 1504
Qy 266 RTGGRKRKILSALYPMIDIQVLSQERQPFHTRCEENGG---CSHLCLLSPSEPFYCA 322
Db 1505 WTGHNVTVQRTNTQPPDLQVYHPSRQPLADNPCEA-NGGKGPCSHLCLINYNRTL-SCA 1562
Qy 323 CPTGVQIQNGRCKCAGAEVILLARTRDLRISLDPDFDIV-LQVDDIRHAIADYD 381
Db 1563 CPHLMKLDKONTTCYE-FKPELLYARQMEIRGVDDNPYNYIISFTVPDIDNVTVDDYD 1621
Qy 382 PLEGVYVWTDDEVARIRRAYLDGSGAOTLVNTEINDPGIADVWVARNLYWT--DTGTR 439
Db 1622 AVEQRIYWSVTRTQIKRAFGINGVETIVSADLPNAHGLSVDWVSRNLFWTSYDTNKKQ 1681
Qy 440 IEVTRLANGTSRKILSVESDLDEPRALALHPVNMGLMYWTDWGNPKIECANLQDQERRVLN 499
Db 1682 INVARLDG-SPEKNAVIOGLDKPHCLVWHPHGLKLYWTD-GDN--ISVANMGSNRTLLFT 1737
Qy 500 ASLQWPNGLALDQEGKLYWGDAXTDKIEVINVDGTRKRTL--LEDKLPILFGFTLLGDF 557
Db 1738 NQRG-PVGLAIDYPEKLYWTSNGNTINRCNLDSGLDEVTVAVKQSLSKATALAIMGDK 1796
Qy 558 IYWTDMQRRSIRVHVVKASRDVILDQLPDLMLGLKAVNVAKV--VGTPNCPADNRGCGSHL 615
Db 1797 LWNADQASERMGTCKKDGTEVTVLRNSTTLVLMKLYVDESIQAGNSPNCVNNGDCQSL 1856
Qy 616 CFFTPHATR-CCPGIGLELLSDMKTCIVPEAPLVFTSRAAHRISLETNN-NDVAIPUTG 673
Db 1857 CLPTSETSRSCWCTAGYSLKSGQSCGVSFLLYSVHGIRGIPLDPNKSDALVPVSG 1916
Qy 674 VKESALDPVSNHHIYWTDVSLKTIISRAFWNGSSVEHVVEFGLDYPSGMADVWGNKLY 733
Db 1917 TSLAVGIDFHAENDIYWDMGLSTISRKRDKQWREDVWTVNGIRGVEGIADVWDTAGNIY 1976

```
QY 734 WADTCTNRIEVARLDGQROVLWRLDLPRLSLALDPTKGYLYTWTGCKRIIVRAFDG 793
Db 1977 WTDQGFVIEVARLNGSRVYVISOGLDKPRAITVHPKGYLFWTEWQYRIERSRLDG 2036
QY 794 TNCMTLVD-KVGRANDLTIDYADQRLYTDLDTNMISSNM-LQGER-VVIADLPHFPFG 850
Db 2037 TERMVLNVISWPNGISVYDEGLKLYWCDAFTKIERIDLETGENREVVLSSDNWDFPS 2096
QY 851 LTQYSDIYIYTDWNLH-STERADKTSGRNRTLIQGHLDLF-VMDILYFHSRQGLNDC 906
Db 2097 VSVFEDYIYMSD-RTHANGSIKRGSKDNATESVSLRTGIGVOLKDIKVENRAKGTNVV 2155
QY 907 MHNNGSQGLCLAIPEGHR-CGCASHYTLDPSSRNCSPPTFLFSPQSASIRMPDPOH 965
Db 2156 AQNNGGCOQLCFRGGGRRTACA-HGLSEDSGVSCRDYDGYLLYSERTILKSHLSDEN 2214
QY 966 SPDLILPL-----HGLRNKAIYDLDPLDKFYWDGRQ-----NIKRAKDDGT- 1008
Db 2215 --NLNAPIKPFEDAHEMKNVIALAFD----YRYGTGSGNRIFYSDIHFGNIQQINDDGTG 2268
QY 1009 QPFVLTSLSQGNPDQRPHDLSID--IYSR---TLFWTCEATNTINVRHLSGEAMG---- 1059
Db 2269 RRTIVENVG-----SVEGLAYHRGWDTLYWTSTTITRTHTVQSRILGAFER 2316
QY 1060 ---VVLGRDRKPRAI VNAERGYLYFTNMODRAAKIERAALDGTREVLFTTGLIRPVA 1116
Db 2317 ETVITMSGD-DHPRAFVLDECONLMFWTNWNEQHSIMRATLSGANVLIIIDQDRTFNG 2375
QY 1117 LVVDNTLKLFWVDADLKRIESCDLSGANRLTLEDANIVQLGLTILGKHLIYWDROOQM 1176
Db 2376 LAIDHRAEKIVFSDATLKDIERCEYDGGSHRVILKSEPHVPGFLAVGYDIYFTWDWVRA 2435
QY 1177 IERVEKTTGDKRTRIQGRVAHL-TGIHAVEVSLSEFSAHPCARDNGCCHICTAKDGT 1235
Db 2436 VORANKYYGTDMKLRIPIPOQPMGIIVAN-DTDSCLSPCRVNNGGCQDLCLLTPKXG 2494
QY 1236 PRCSFVHLVLLQNLTLTCEPPTCS-PQOFACATGEIDICPGAWRCDFGPBCDQSDDE- 1293
Db 2495 VNCSCRGRLVQEDFTCKALNSTCNVHDEFGNG--DCIDFSRTCDGVVHCKKSDSKQ 2552
QY 1294 ---GCFVCSAAQFCPCARQCQVDRLRCDGEADCDQDRSDEADC-ALICLPNFRCASGQCV 1349
Db 2553 SYCSRSKCKGFLHCHMGRCVASFRCWGVNDGCGNSDEVFCNKTSCTAATEFRCDGSCI 2612
QY 1350 LIKQCDSPFCIDGSDLMCEITKPPSDSPSAHSAIGPVIGIILSLFVWGGVYFVQCR 1409
Db 2613 GNSGRCNQFIDCEDASDEMNCTATDCSS-----YFKLGVKGTTFQKCEH 2656
QY 1410 V-----VQORVAGAN--GPFPEHY---VSGTTPHVLPLNFIA-PGG 1442
Db 2657 TSLCVAPSWVCD--GANDCGDYSERNCPGGRKPKCPANFYACPSG 2700
RESULT 7
alpha-2-macroglobulin receptor precursor - mouse
N:Alternate names: CP91; LDL receptor-related protein 1; low density lipoprotein receptor
A:Species: Mus musculus (house mouse)
C:Date: 04-Sep-1998 #sequence_revision 04-Sep-1998 #text_change 09-Jul-2004
C:Accession: S25111; S32554
R:van Leuven, F.
submitted to the EMBL Data Library, July 1992
A:Reference number: S25111
A:Accession: S25111
A:Molecule type: mRNA
A:Residues: 1-4545 <VANI>
A:Cross-references: UNIPROT:Q61291; EMBL:X67469; NID:g49941; PIDN:CAA47817.1; PID:g49942
R:van Leuven, F.; Stas, L.; Raymakers, L.; Overbergh, L.; de Strooper, B.; Hilliker, C.;
Biochim. Biophys. Acta 1173, 71-74, 1993
A:Title: Molecular cloning and sequencing of the murine alpha-2-macroglobulin receptor c
A:Reference number: S32554; MUID:93250049; PMID:8485155
A:Accession: S32554
A>Status: nucleic acid sequence not shown
```

```
A:Molecule type: mRNA
A:Residues: 1-28;4416-4453 <VANI>
A:Cross-references: EMBL:X67469
C:Comment: The alpha-2-macroglobulin receptor complex consists of noncovalently-associated
ciated protein (see PIR:J0X0281).
C:Superfamily: alpha-2-macroglobulin receptor; EGF homology; LDL receptor ligand-binding
C:Keywords: beta-hydroxyasparagine; beta-hydroxyaspartic acid; calcium binding; glycoprotein
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-3944,3945-4545/Product: alpha-2-macroglobulin receptor #status predicted <MAT>
F:20-3944/Domain: alpha-2-macroglobulin receptor 515K chain #status predicted <515K>
F:28-65/Domain: LDL receptor ligand-binding repeat homology <LDL1>
F:73-109/Domain: LDL receptor ligand-binding repeat homology <LDL2>
F:116-149/Domain: EGF homology <EG1>
F:155-189/Domain: EGF homology <EG2>
F:199-240/Domain: LDL receptor YWTD-containing repeat homology <YW01>
F:241-282/Domain: LDL receptor YWTD-containing repeat homology <YW02>
F:293-335/Domain: LDL receptor YWTD-containing repeat homology <YW03>
F:336-379/Domain: LDL receptor YWTD-containing repeat homology <YW04>
F:380-421/Domain: LDL receptor YWTD-containing repeat homology <YW05>
F:422-469/Domain: LDL receptor YWTD-containing repeat homology <YW06>
F:479-520/Domain: EGF homology <EG3>
F:572-614/Domain: LDL receptor YWTD-containing repeat homology <YW07>
F:615-660/Domain: LDL receptor YWTD-containing repeat homology <YW08>
F:661-711/Domain: LDL receptor YWTD-containing repeat homology <YW09>
F:712-753/Domain: LDL receptor YWTD-containing repeat homology <YW10>
F:754-800/Domain: LDL receptor YWTD-containing repeat homology <YW11>
F:808-843/Domain: EGF homology <EG4>
F:855-891/Domain: LDL receptor ligand-binding repeat homology <LDL3>
F:896-932/Domain: LDL receptor ligand-binding repeat homology <LDL4>
F:937-972/Domain: LDL receptor ligand-binding repeat homology <LDL5>
F:977-1012/Domain: LDL receptor ligand-binding repeat homology <LDL6>
F:1016-1052/Domain: LDL receptor ligand-binding repeat homology <LDL7>
F:1063-1098/Domain: LDL receptor ligand-binding repeat homology <LDL8>
F:1105-1141/Domain: LDL receptor ligand-binding repeat homology <LDL9>
F:1146-1183/Domain: LDL receptor ligand-binding repeat homology <LDL10>
F:1186-1222/Domain: EGF homology <EG5>
F:1228-1262/Domain: EGF homology <EG6>
F:1270-1309/Domain: LDL receptor YWTD-containing repeat homology <YW12>
F:1310-1356/Domain: LDL receptor YWTD-containing repeat homology <YW13>
F:1357-1399/Domain: LDL receptor YWTD-containing repeat homology <YW14>
F:1400-1446/Domain: LDL receptor YWTD-containing repeat homology <YW15>
F:1447-1489/Domain: LDL receptor YWTD-containing repeat homology <YW16>
F:1490-1532/Domain: LDL receptor YWTD-containing repeat homology <YW17>
F:1541-1579/Domain: EGF homology <EG7>
F:1584-1627/Domain: LDL receptor YWTD-containing repeat homology <YW18>
F:1628-1670/Domain: LDL receptor YWTD-containing repeat homology <YW19>
F:1671-1714/Domain: LDL receptor YWTD-containing repeat homology <YW20>
F:1715-1754/Domain: LDL receptor YWTD-containing repeat homology <YW21>
F:1755-1797/Domain: LDL receptor YWTD-containing repeat homology <YW22>
F:1798-1847/Domain: LDL receptor YWTD-containing repeat homology <YW23>
F:1851-1887/Domain: EGF homology <EG8>
F:1935-1977/Domain: LDL receptor YWTD-containing repeat homology <YW24>
F:1978-2020/Domain: LDL receptor YWTD-containing repeat homology <YW25>
F:2021-2064/Domain: LDL receptor YWTD-containing repeat homology <YW26>
F:2065-2106/Domain: LDL receptor YWTD-containing repeat homology <YW27>
F:2107-2152/Domain: LDL receptor YWTD-containing repeat homology <YW28>
F:2160-2195/Domain: EGF homology <EG9>
F:2200-2242/Domain: LDL receptor YWTD-containing repeat homology <YW29>
F:2254-2295/Domain: LDL receptor YWTD-containing repeat homology <YW30>
F:2345-2389/Domain: LDL receptor YWTD-containing repeat homology <YW31>
F:2390-2430/Domain: LDL receptor YWTD-containing repeat homology <YW32>
F:2431-2474/Domain: LDL receptor YWTD-containing repeat homology <YW33>
F:2483-2518/Domain: EGF homology <EG10>
F:2525-2562/Domain: LDL receptor ligand-binding repeat homology <LDL11>
F:2567-2601/Domain: LDL receptor ligand-binding repeat homology <LDL12>
F:2606-2640/Domain: LDL receptor ligand-binding repeat homology <LDL13>
F:2653-2689/Domain: LDL receptor ligand-binding repeat homology <LDL14>
F:2697-2731/Domain: LDL receptor ligand-binding repeat homology <LDL15>
F:2735-2770/Domain: LDL receptor ligand-binding repeat homology <LDL16>
F:2775-2813/Domain: LDL receptor ligand-binding repeat homology <LDL17>
F:2819-2854/Domain: LDL receptor ligand-binding repeat homology <LDL18>
F:2859-2898/Domain: LDL receptor ligand-binding repeat homology <LDL19>
F:2905-2940/Domain: LDL receptor ligand-binding repeat homology <LDL20>
```

F:2945-2981/Domain: EGF homology <EG11>
F:2987-3012/Domain: EGF homology <EG12>
F:3030-3069/Domain: LDL receptor YWTD-containing repeat homology <YW34>
F:3070-3114/Domain: LDL receptor YWTD-containing repeat homology <YW35>
F:3115-3157/Domain: LDL receptor YWTD-containing repeat homology <YW36>
F:3158-3201/Domain: LDL receptor YWTD-containing repeat homology <YW37>
F:3202-3242/Domain: LDL receptor YWTD-containing repeat homology <YW38>
F:3243-3285/Domain: LDL receptor YWTD-containing repeat homology <YW39>
F:3295-3331/Domain: EGF homology <EG13>
F:3335-3370/Domain: LDL receptor ligand-binding repeat homology <LDL1>
F:3375-3409/Domain: LDL receptor ligand-binding repeat homology <LDL2>
F:3414-3449/Domain: LDL receptor ligand-binding repeat homology <LDL3>
F:3454-3490/Domain: LDL receptor ligand-binding repeat homology <LDL4>
F:3495-3532/Domain: LDL receptor ligand-binding repeat homology <LDL5>
F:3537-3571/Domain: LDL receptor ligand-binding repeat homology <LDL6>
F:3576-3610/Domain: LDL receptor ligand-binding repeat homology <LDL7>
F:3614-3648/Domain: LDL receptor ligand-binding repeat homology <LDL8>
F:3655-3691/Domain: LDL receptor ligand-binding repeat homology <LDL9>
F:3696-3732/Domain: LDL receptor ligand-binding repeat homology <LDL10>
F:3742-3777/Domain: LDL receptor ligand-binding repeat homology <LDL11>
F:3786-3823/Domain: EGF homology <EG14>
F:3829-3861/Domain: EGF homology <EG15>
F:3869-3912/Domain: LDL receptor YWTD-containing repeat homology <YW40>
F:3913-3970/Domain: LDL receptor YWTD-containing repeat homology <YW41>
F:3945-4545/Domain: alpha-2-macroglobulin receptor 85K chain #status predicted <85K>
F:3945-4421/Domain: 85K chain extracellular #status predicted <EXT>
F:3971-4013/Domain: LDL receptor YWTD-containing repeat homology <YW42>
F:4014-4057/Domain: LDL receptor YWTD-containing repeat homology <YW43>
F:4058-4100/Domain: LDL receptor YWTD-containing repeat homology <YW44>
F:4101-4143/Domain: LDL receptor YWTD-containing repeat homology <YW45>
F:4152-4183/Domain: EGF homology <EG16>
F:4201-4232/Domain: EGF homology <EG17>
F:4237-4268/Domain: EGF homology <EG18>
F:4273-4304/Domain: EGF homology <EG19>
F:4309-4340/Domain: EGF homology <EG20>
F:4345-4375/Domain: EGF homology <EG21>
F:4378-4409/Domain: EGF homology <EG22>
F:4422-4445/Domain: transmembrane #status predicted <TM>
F:4446-4545/Domain: intracellular #status predicted <INT>
F:167,2999/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted
F:2959/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status predicted
F:4076,4126,4279/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 24.2%; Score 2113; DB 1; Length 4545;
Best Local Similarity 34.3%; Pred. No. 3.5e-124;
Matches 508; Conservative 265; Mismatches 588; Indels 122; Gaps 51;

Qy 35 LLFANRRDRLVDAGGVKLESTIVVSGLEDAADVDFQFSGAVYWTVDSEBAIKQTYLNQ 94
Db 1272 IIFSNRHEIRIDLH--KGDYSVLVPLGRLNTIALDFHLSQSALYWTDAVEDKIYRGKLLD 1329

Qy 95 TGA--AVQNVVISGLVSPDGLACDWGKKLYWTDSETNRIEVANLNGTSRKVLFWQDLQ 152
Db 1330 NGALTSFEWIIQYGLATPEGLAWDIAAGNIYVWESNLDAQIEVAKLDGTLRTLAGDIEH 1389

Qy 153 PRALADPAHGYMYWTOW-GSTPRIERAGMDGSTRKII--VDSDIYWPNGLTIDLEROKL 209
Db 1390 PRALADPRGILFWTDASLPRIEASMSGARRTIHRTGSGGCANGLTVDYLEKRI 1449

Qy 210 YWADAKLSFIHRANLDGSRFQKVEGS--LTHPFAITLSGDTLYWTDQTRSIIHACNKR 267
Db 1450 LWIDARSDAIVSARYDGSGHMEVLRGHEFLSHPPFAVTLYGGEVYWTDWRTNTLAKANKWT 1509

Qy 268 GKKRKEILSALYSPMDQVLVSQRPQPHTRCEBDNGG---CSHLCLLSPSEPPYTCACP 324
Db 1510 GHNVTVQRTNTQFDLQVHPSPQMAPNPCEA-NGGRGPCSHLCLINVRT--VSWACP 1567

Qy 325 TGVOLONGRTCKAGAEVILLARRTLRRISLTPDFTDIV-LQVDDIRHAIADYDPL 383
Db 1568 HMKLHKDNTTCYB-FKKFLYAKMEIRGVLDAPYNYIISTVPDINDVTVLDYDAR 1626

Qy 384 EGYVYVTDDEVAIRRAYLDGSGAQTLVNTBNDPDGIADVDAWARNLYWT--DTGTDRIE 441
Db 1668 YAPSWVCD---GANDCGDYSDERDCPGVGRPCPLNYFACPSG 2707

Db 1627 EQRVYWSVDRTOAIKRAFINGTGVETVVVSADLPNAHGLAVDWVSRNLFWTSDYTNKKQIN 1686

Qy 442 VTRLINGTSRKTLVSEDLDEPAIALHPVMGLMYKMTDNGENPKIECANLDDGQERRVLVNAS 501
Db 1687 VARLDGSGFKNAV--QGLEQPHGLVHPURGLKLYWTD-GDN--ISNANMDGSHNTLLFSGQ 1742

Qy 502 LGWPNGLALDLQEGKLGWDAKTDKIEVINVDGTYKRRTL--LEDKLPHFIFGFTLLGDPIY 559
Db 1743 KG-PVGLAIDPESKLYWISSGNHTINRCNLDDGSELEVIDTMRSQLGKATALAINGDKLW 1801

Qy 560 WTDQWRSIERVHKVKSARDVIIDQLPDL-MGLKAVNVA---KVVGTPNCPADNRNGCCSHL 615
Db 1802 WADQVSEKMGTCNKADGSGSVLRNRTTLVNMKVDYDESIQLEHEGTTPCPSVNNGDCSQL 1861

Qy 616 CFFTPHATR-CGCPGLELLSDMKTICVPEAFVFTSPAAIHRISLEWNN-NDVAIPLTG 673
Db 1862 CLPTSETTRSCMTAGYSLRSGQACGEGVSGFLYSVHEGIRGIPLDPNDSKDALVPVSG 1921

Qy 674 VKEASALDFDVSNHHIYWTVDVSLKTIISRAFNGSGSVHEVVEFGLDYPEGMAVDMGKNLY 733
Db 1922 TSLAVGIDFHAENDTIYWDVMDGLSTISRAKEDQIWRDVTNGIRVSGIAVDWTAGNIY 1981

Qy 734 WADGTNRIRIEVARLDGQFQVLRVWRDLNPRSLALDPTKGYIYWTGCKPRIVAFMDG 793
Db 1982 WTDQGFVIEVARLNGSFYVVISQGLDKPRAITVHPPEKGYLFTWEGHYPRIERSRLDG 2041

Qy 794 TNCMTLYD-KVGRANDLTIDYADORLYWTDLDTNMISSNM-LQOER-VVITADDLPHRPF 850
Db 2042 TERVVLVNVSVISWPNGISVDYQGGKLYWCARDMDKIERIDLETGENREVLSNNMDMFS 2101

Qy 851 LTOYSDIYIYWDVMDLH---SIERADTKTSGRNRITLQGHLDL-VMDILVFHSRODGLNDC 906
Db 2102 VSPVEDIYMSD-RTHANGSIKRCCKDNATDSVPLRTGIGVQLKDIKVFNRDRQKGTNV 2160

Qy 907 MHNGGCGQLCLATPGGHR-CGCASHYTLDPSSRNCSPPTFTLLFSQKSAISRM-IPDDQ 964
Db 2161 AVANGGCGQLCLYRGGGQACACA-HGMLAEDGASCREYAGYLLYSERTILKSHLSDER 2219

Qy 965 HSPDLIILPLHG---LRNVKAIYD-----PLDKFYIYWDGR-ONTKRAKDDGTQPFVL 1013
Db 2220 NLNAPQFDFDEPHMKVIALAFDYRAGTSPGTNRIFFSDIHFGNIQOINDGSG---R 2276

Qy 1014 TSLSQGQNPDRQPHDLSD--IYSR---TLFWTCEATNTINVHRLSGEAMG-----VW 1061
Db 2277 TTIENVVG-----SVEGLAYHGWDTLYWTSYTTSTIHTVTDQTRPGAFERETVIT 2328

Qy 1062 LRGRDRKPRIVNAERGYLYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVVALVDN 1121
Db 2329 MSGD-DHPRAFVLDECQNLMTNBNLHPSIMRAALSGANVLTLIEKDIRTPNGLAIDH 2387

Qy 1122 TLGKLFWDADLKRIESCDLSGANRLTLEDANIYQPLGLTILGKLYWIDRQOQMIERVE 1181
Db 2388 RAELKLYPSDATLDKIERCEVDGSHRYVILKSEPVHPFLAVYGEHIFMTDWRRAVQRAN 2447

Qy 1182 KTTGD--KRTRIOGRVAHLTGIIHAVERVSELEFSAHPCARDNGGCSHCIAKGGDTPRCS 1239
Db 2448 KYVGSDMKLLRVDPQQPMGLIAVANDTNSCLS--PCRINNGCGQDLCLLTHQGVNCS 2505

Qy 1240 CPVHLVLLQNLITCGEPPTC-SPDQFACATGEIDCIPGANRCDDGFPEDCDSDSE---G 1294
Db 2506 CRGRIQLQEDFTCRVNSSCRAQDEFECANGE--CISFSLTCDGVSHCKKSDKPSYCN 2563

Qy 1295 CPVCSAAQFCARGQCYDLRLRCDEADCDQRSDEADC-AICLPNQFRCSGQCVLKQ 1353
Db 2564 SRCKKTFRCQNGRCVSNMLWCNGVDYCGDGSDEIPCNKTACGVBFRCDGSCIGNSS 2623

Qy 1354 QCSFPDIDGSDBLMCEITKPPSDSDPSAHSAIGPVIGITLSLFWMGVYFVQV--- 1410
Db 2624 RCNOFVDCEDASDEMNCSATDCSS-----YFLGVKGVLPQPCERTSLC 2667

Qy 1411 -----VCORVAGAN--GPFPEHY-VSGT--PHVPLNFIA-PGG 1442
Db 2668 YAPSWVCD---GANDCGDYSDERDCPGVGRPCPLNYFACPSG 2707


```
QY 1218 ARDNGCSCIACGD-----GTRCSC-----PVLVLQN----- 1249
Db 2733 KTDNGGCDLCTVADDIGLAASKVQCSCNDTYELVFPKGDYPTQCVRGNSSPAKEC 2792
QY 1250 -----LL--TCGPPTCS-----PDQFACATGEIDCIPGAW 1278
Db 2793 LPPYFQCQDGCSCILLGATCDKPCADASDENPNYCNTRSCPENYLCNTR-RCIDSAK 2851
QY 1279 RCDGFPECDQSDREGCP---VCSAAQPPCARGQCVRDLRLRCDGBADCCDR--SDEA--- 1330
Db 2852 KCHNIDDCGDSDELDCPSAVACABGTFCPSNGHCINQTKVCDGHNDCHDEQVSDSLAT 2911
QY 1331 -----DCDAI-----CLPNQPRCASG 1346
Db 2912 CPGLPIDCRGVKRCFNNNICIQPADLDCGVDDCGKADENQLFCMQNCCQAQHYVRCPSG 2971
QY 1347 QCVLKQCCDSPDCIDGSDLMCEITKPPSDSPAHSSAIGPVGIGIILSLFVMGVVFPV 1406
Db 2972 RCIPETWCQDGDNDSCDGDWETHNCT-----DTAGKKICVGDYL-----FQ 3013
QY 1407 QRVVVC 1412
Db 3014 CDNLAK 3019

RESULT 10
T46336
Hypothetical protein DKFZp43400213.1 - human
C:Species: Homo sapiens (man)
C>Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
C:Accession: T46336
R:Koehret, X.; Beyer, A.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A:Reference number: Z23037
A:Accession: T46336
A>Status: preliminary
A.Molecule type: mRNA
A.Residues: 1-252 <ANA>
A.Cross-references: UNIPROT:O9NSY4; EMBL:AL137651
A.Experimental source: adult testis; clone DKFZp43400213
C:Genetics:
A.Note: DKFZp43400213.1
C:Superfamily: low density lipoprotein receptor-related protein; EGF homology; LDL recep

Query Match 11.7%; Score 1024; DB 2; Length 252;
Best Local Similarity 83.5%; Pred. No. 2.8e-57;
Matches 193; Conservative 16; Mismatches 22; Indels 0; Gaps 0;
QY 383 LEGVYVWTDDEVRAIRRAYLDGSGAOTLVNTEINDPGIADVWVARNLYTDTGTDRTEV 442
Db 1 MEGHYWTDDEWAIIRRAYLDGSGAOTLINTKINDPDDIAVNWARSLYTHTGTEHIEV 60
QY 443 TRLNGTSKILVSEDLDEPRAIALHPVGMGLMYTWDWGNPKIECANLDGQERRVNVNASL 502
Db 61 TCLNSTSHKILVSEDMDEPRAIALHPENGLYTWIDWGNPSEIKRANLDRQELRVNASL 120
QY 503 GWPNGALDLQEGKLYGDAKTDEKIEVINVDGTRKRTLEKLPFHIFGFTLLGDFIYWD 562
Db 121 GWPNGALDLQEGKLYGDAKTDEKIEAISVDETRKQTLKDLKLPFHIFRFTLLGDFIYNTA 180
QY 563 WQRSIERVHVKASRDVIIIDQLPDLMLKAVNVKVVGTNPNPCADRNGGCS 613
Db 181 WQHSIERVHVKANRVDIIIDQLPDLMLKAVNVKVVGTNPNPHADRNGGAA 231

RESULT 11
T18308
probable vitellogenin receptor - yellow fever mosquito
C:Species: Aedes aegypti (yellow fever mosquito)
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T18308
R:Sappington, T.W.; Kokoza, V.A.; Cho, W.L.; Raikhel, A.S.
```

```
Proc. Natl. Acad. Sci. U.S.A. 93, 8934-8939, 1996
A>Title: Molecular characterization of the mosquito vitellogenin receptor reveals unexpect
A:Reference number: Z18874; MUID:96392343; PMID:8799131
A:Accession: T18308
A>Status: preliminary; translated from GB/EMBL/DBJ
A.Molecule type: mRNA
A.Residues: 1-1847 <SAP>
A.Cross-references: UNIPROT:O76952; EMBL:L77800; NID:g3386582; PID:g3386583; PIDN:AAC2845
C:Function:
A.Description: involved in endocytosis
F:45-79/Domain: LDL receptor ligand-binding repeat homology <LDL1>
F:84-120/Domain: LDL receptor ligand-binding repeat homology <LDL2>
F:132-167/Domain: LDL receptor ligand-binding repeat homology <LDL3>
F:174-207/Domain: LDL receptor ligand-binding repeat homology <LDL4>
F:211-249/Domain: LDL receptor ligand-binding repeat homology <LDL5>
F:256-297/Domain: LDL receptor ligand-binding repeat homology <LDL6>
F:298-333/Domain: LDL receptor ligand-binding repeat homology <LDL7>
F:334-369/Domain: LDL receptor ligand-binding repeat homology <LDL8>
F:370-405/Domain: LDL receptor ligand-binding repeat homology <LDL9>
F:406-441/Domain: LDL receptor ligand-binding repeat homology <LDL10>
F:442-477/Domain: LDL receptor ligand-binding repeat homology <LDL11>
F:478-513/Domain: LDL receptor ligand-binding repeat homology <LDL12>
F:514-549/Domain: LDL receptor ligand-binding repeat homology <LDL13>
Query Match 10.5%; Score 914.5; DB 2; Length 1847;
Best Local Similarity 24.9%; Pred. No. 4.1e-49;
Matches 299; Conservative 150; Mismatches 386; Indels 365; Gaps 46;
QY 299 CEEDN-----GGCSHLCLLSPSEPFYTCACPTGVQLQONGRTCKA--GAEEVLLIARRTD 351
Db 287 CEDINECECYGLCSQGCENTPGS--FKCTCVDFKFKLDDSRTECLDSDSTELLTYTQKS 344
QY 352 LRRISLTPDPTDVLQVDDIRHAIADYDPLEGYVYVWD--DEVRAIRRAYLDGSGAQT 409
Db 345 IGGHLNLTQHYV--AKDLSQVIGSYD--GRHVYVTDISPFKTESLSDGTREL 399
QY 410 LVNTEINDPGIADVWVARNLYTDTGTDRTEVRLNGTSKILVSEDLDEPRAIALHPV 469
Db 400 LITSGLASPEDELDWLTGNIYFSDSGHMMIAVCSNNGVHCTILIQTLHPRGIALMPQ 459
QY 470 MGLMYTWDWGNPKIECANLDGQERRVNVNASLGNLWPNGLALDLQEGKLYGDAKTDEKIEV 529
Db 460 NGTLFYSYDWDGNAMIGAAGDGNKRLIEQDIHWPNGCLDWPNGRLYVWDKXKIES 519
QY 530 INVDTGTRTLEKLPFHIFGFTLLGDFIYWDWQRSIERVHVKVA-SRDVIID--QLP 586
Db 520 IRVDGTRVTVLADVLKHPFSIAVFNDRLYWSDWDTKSIQSCDKFNGKDKRIVVHDKRIF 579
QY 587 DL-----MGLKAVNVKVVGTNPNPCADRNGGCSHLCTFFTPH-ATRCGCPIGLELLDMKTC- 640
Db 580 DVHIYHSGLP-----KGDPCL--GTFCSHLCLLAPNDSYSCACPYGMSLKADKHSR 631
QY 641 -IVEAPFLVFTSRAAHRIISLETNNNDVAIPLTGKVASALD-----FVSNHHI 689
Db 632 ETVKRQYLLV--GIANYLVLTETQ-----FGRHSSQADAYQIFFRMAFNSITGEI 682
QY 690 YWTDVSLKTSIRAPMGSSVHVVEFGLDYPEGVADVWGMKNLYWADTGNIIRIVARDLG 749
Db 683 FVADNRQAKFTVDPDKTKSSQKLITTTGIGNISALAFDGLNLYWTDSEKSTVEVFSLOT 742
QY 750 QFRVLY-WEDLONPRSLALDPTKGIY-----WTEWGGKPRIVRAF 790
Db 743 RHRAILQHYLGQDIPVGLAIVSEGMKFIARSLPVPVPTHIDRLDWTGRGPHLVIEER 802
QY 791 MDGTCMTLVKGRANDLTIDYADQLYWTDLDTNNIESNNMLGQERRVVIADLPHFPFG 850
Db 803 LSGNGSFNV-----IDRLRTVYVWDMGSSKIEFTSYEGDTRHLFRFLRPLVS 852
QY 851 LTQVSDVIYVWDMNLSHIERADKTS--GRNRTLQGHLD-----FVMDILVHFS--SRQ 900
Db 853 IAVGDSIFWTCRYSKRLYWSDKHNLGVTKKIT---IDKPPYGAFPPDEIVLLGSGQPLQRY 909
QY 901 DGLNDCMHNNGCQCQLCLAIPEGHRCGASHYTLDPSSRNCSPPTTFLFSQKSAISRM 960
```

```

Db      910 D--HPCMKQNGCCHIC--VPAG-----MYSSACICPTGMIFFSPK----- 946
Qy      961 PDDQHSPLILPLHLNRVAIDYDPLDKFYIYWDGRQNIKRAKDGTQPFVLTSLSQGG 1020
Db      947 -----NTTCID----- 952
Qy      1021 NPDRQPHLSIDISYRTLFWTCEATNTINVHRLSGEAMGVVLR--GDRKPRAIIVNAER 1078
Db      953 -----AID-----CEPKCT-----SGECLTISKRCNGKD----- 977
Qy      1079 GYLFTNMQDRAAKIERAALDGTREVLFTTGLIRPVVALVVDNTLGLFWVDADLKRIES 1138
Db      978 -----CADGSD----- 986
Qy      1139 CDLSGANRLTLEDANIVQPLGLTLGKHLVY-----IDRQOQMIHVEKTTGDK 1187
Db      987 CDEAG-----QPKQL-----HCQYDEFMCADKSKCIDQTRCDEHVDGSD 1029
Qy      1188 RTRIQRVAHLTGTHAVEEVSLEBFSAPCARDNGGC---SHICIAKGDGTPRC---SCP 1241
Db      1030 EMKEG-YDRGTGCH-----EHOACPGMGIDVNTLC-----DGFPDCLDGSDE 1073
Qy      1242 VHLVLLQNLITCGEPPTCSPDPQACATGEIDCIPGAWRCDFPDCDQSDSDE-EGCPV--- 1297
Db      1074 VGCTDLTN--EKSNAITCGPLMPCNMGQ--CIPKWECDGNPDCTDGSDEHDKLTKTD 1129
Qy      1298 CSAAQFPCARGQCVDLRLRCDEADCDRDEADCD----- 1333
Db      1130 CGAGFTKALGHCHIEDRLCLDGNNDGNSDELNCKVLEBFCVGLGEDDNTKYLCPRSKG 1189
Qy      1334 -----AICLPNPRCASGCVLKIQQCKQDQSFPCDIDGSDLMC 1370
Db      1190 CLDIAVRNGTAECPCDEDEAGSCNGLQBFQCKSGKIRKEWRCKEVDCCDGSDEVC 1249

RESULT 12
T13171
probable vitellogenin receptor - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004
C;Accession: T13171
R;Schonbaum, C.P.; Lee, S.; Mahowald, A.P.
Proc. Natl. Acad. Sci. U.S.A. 92, 1485-1489, 1995
A;Title: The Drosophila yolkless gene encodes a vitellogenin receptor belonging to the 1
A;Reference number: 217627; MUID:95183490; PMID:7878005
A;Accession: T13171
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1984 <SCH>
A;Cross-references: UNIPROT:P98163; EMBL:U13637; NID:G535345; PID:G535346; PIDN:AAB60217
C;Genetics:
A;Gene: yl
A;Cross-references: FlyBase:FBgn0004649
A;Map position: 1
P;90-124/Domain: LDL receptor ligand-binding repeat homology <LDL1>
P;129-166/Domain: LDL receptor ligand-binding repeat homology <LDL8>
P;184-220/Domain: LDL receptor ligand-binding repeat homology <LDL10>
P;227-262/Domain: LDL receptor ligand-binding repeat homology <LDL11>
P;266-304/Domain: LDL receptor ligand-binding repeat homology <LDL12>
P;1025-1062/Domain: LDL receptor ligand-binding repeat homology <LDL11>
P;1074-1109/Domain: LDL receptor ligand-binding repeat homology <LDL2>
P;1118-1152/Domain: LDL receptor ligand-binding repeat homology <LDL3>
P;1158-1193/Domain: LDL receptor ligand-binding repeat homology <LDL4>
P;1198-1232/Domain: LDL receptor ligand-binding repeat homology <LDL5>
P;1243-1279/Domain: LDL receptor ligand-binding repeat homology <LDL6>
P;1283-1318/Domain: LDL receptor ligand-binding repeat homology <LDL7>
P;1340-1375/Domain: LDL receptor ligand-binding repeat homology <LDL9>

Query Match      10.5%; Score 913.5; DB 2; Length 1984;
Best Local Similarity 23.8%; Pred. No. 5.3e-49;
Matches 305; Conservative 162; Mismatches 434; Indels 381; Gaps 47;

```

```

Qy      299 CEEDNGGCSHLCLLSPSEPPYTCACPTGVOLQNGRTCKA---GABE---VLLIARRTDL 352
Db      352 CKQDDLLCSQGC--ENTSGGYRCVCDAGYLLDKDNRITCRAVYVYSGKEQQPLLLLTTQWTI 409
Qy      353 RRISLDPDFTDIVLQV--DDIRHAIAIDYDPLEGYVYWD--DEVRAIRRAYLDGSGAQT 409
Db      410 MGHRLREDNRNHYVQVAGNLSKVIGVAYD--GSHIYWTNIQNEABESIVKANGDGSNAEI 467
Qy      410 LVNTEINDPGIADVARNLYWTDCTGTDRIVETRLNGTSRKILVSDLEDPRAIALHPV 469
Db      468 LLTSGLDAPEDLAVDLWTQNIYFSDNIMRHIAVCSNDGLNCAVLTVQDVHQPRLSAWVPQ 527
Qy      470 MGLMYWTDWENPKIECANLDGERRVLVNASLGPNGLALDLOEGKLYMGDAKTDKIEV 529
Db      528 KGLMFWTDGKPMIGRASMDGSRSPVSDNIEMWPNGLDMHQRIYWDYDAKLSVQT 587
Qy      530 INVDTGTRTLLEDKLPHIFGFTLLGDFIYWTDMQRSIERVHKVKASRDVILIDPLDM 589
Db      588 VRPDGTGRTRVLDMGLKHPYGLAIFEDQLYMSDWATKSVHACHKFSKGDHRIILAKDRIT 647
Qy      590 GLKAVNVKVVGTNPACDRNGGCSHLCTFFTP-----HATRCGCPGLELLSDMTCTIYPE 644
Db      648 AVHIYHPAK--OPNSPHGCENATCSHLCLLAEPEIGGHS--CACPDGMRLAPDHRRCMLME 704
Qy      645 ----AFLVFTSRAAIHRISLETNNNDVAIPLTVKKEAS-----ALDPDVSNNNHIYTDV 694
Db      705 KRQLFI-----GLGVLLIEHT-----AFGAHVQSKSYTLPCLINEWYNNRINGSLI 753
Qy      695 SLKTIISRAFWNGSSVEH-----VVEFGLDYPEGMAVDMGNKLYWADTGTNRIVEARLDGQ 750
Db      754 IADNQRLILEFQEPESHESNVLVRSNLGNVSALAFDHLRNLRYWADTERAVIEVLSLQTR 813
Qy      751 FRQVL--VWRDLNPRSLALDPTKGIYWT-----EWGKPRIVRAFWD--GTN 795
Db      814 HRALIRFPQGVPIGLTVMPAEGYLVVILKAKRHSIDKIPISGKGEQVHFDEDDLGD 873
Qy      796 CMTLVKVGGRANDLITIDYADQRLYWTDLDTNMISSNNMGLQERVIADLPHFPGLTOYS 855
Db      874 DIKLV-----TDYETQIFWSDSLGRISYVNRVPHSQIFRGKLRPRPSLAMVH 923
Qy      856 DYIYWTDMNLHSIERADTKTSGRNRTLIOGHLDLFWD-----ILVFH 896
Db      924 HDLFWNELGTPIRYWTHKSNMGRPKVID----IMEKDDPAAIMPYPVPVATPNGIPLAAS 978
Qy      897 SSRQDGLDCHWNQCGQLCLAIPEGHR--CGCASHYTL--DPSSRNCSPPTFLFLFSQS 954
Db      979 SPVQESHPCQQNGGCSHCICVGEPPYHSLCPAGFVYRDAGNRTC----- 1025
Qy      955 AISRMIPDDQ--HSPDLILPLH---GLRNVKAIDYDPLDKFYIYWDGRQNIKRAKDDGTQP 1010
Db      1026 -VEALDCEFRCHSGECLTMNRCNGRDC-----VDNSDEMNCDEHRHKP 1070
Qy      1011 FVLTSLSQGNQPNDRPHDLSDIYSRTLFWTCEATNTINVHRLSGEAMGVVLRGDRDKPR 1070
Db      1071 KVLCSPSQ-----FACH-----SSEQ----- 1086
Qy      1071 AIVUNAERYLYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVVALVVDNTLGLFWVD 1130
Db      1087 --CVDKER----- 1092
Qy      1131 ADLKRIESCDSGANRLTLDANIVQPLGLTLGKHLIYWDROOQMIHVEKTTGDKRTR 1190
Db      1093 -----RCD---NRKCDHDS-----DQHCKEPKDS----- 1115
Qy      1191 IQGRVAHLTGTHAVEEVSLEBFSAPCARDNGGC---SHICIAKGDGTPRCSPVHLVLL 1247
Db      1116 ---KKCH---VH-----OHGC--DNGKCVDSLSVC---DGTNDGCDN----- 1146
Qy      1248 QNLLTCEPPTCSPDQACATGEIDCIPGAWRCDFPDCDQSDSDEGCPV---CSAAQFP 1304
Db      1147 SDELLCEATLRCFPMFQCGSG--SCIAGSWECDGRIDCSGSDHDKCVHRSPDPMHR 1204
Qy      1305 CARQCVDLRLRCDEADCDRDEADCD-----ICLPN----- 1339

```


QY 917 CLAIRPGHRCG 928
 Db 1334 CLLRPGGYSCAC 1345

RESULT 14
 EGHU

N:Altermate names: urogastone precursor - human
 C:Species: Homo sapiens (man)
 C:Date: 30-Nov-1980 #sequence revision 14-Aug-1998 #text change 09-Jul-2004
 C:Accession: A25531; A33517; A29721; S45282; S45283
 R:Bell, G.I.; Fong, N.M.; Stempelen, M.M.; Wormsted, M.A.; Caput, D.; Ku, L.; Urdea, M.S.
 Nucleic Acids Res. 14, 8427-8446, 1986
 A:Title: Human epidermal growth factor precursor: cDNA sequence, expression in vitro and
 A:Reference number: A25531; MUID:87066721; PMID:3491360
 A:Accession: A25531
 A:Molecule type: mRNA
 A:Residues: 1-1207 <BBL>
 A:Cross-references: UNIPROT:P01133; EMBL:X04571; NID:g31120; PIDN:CRAA28240.1; PID:g31121
 A:Note: 708-Met was also found
 A:Note: intron positions were also determined
 R:Gregory, H.; Preston, B.M.
 Int. J. Pept. Protein Res. 9, 107-118, 1977
 A:Title: The primary structure of human urogastone.
 A:Reference number: A01388; MUID:77117897; PMID:300079
 A:Accession: A01388
 A:Molecule type: protein
 A:Residues: 971-1023 <GRE>
 A:Note: some of the molecules lack Arg-1023
 R:Furuya, M.; Akaishi, S.; Hirayama, K.
 Biochem. Biophys. Res. Commun. 163, 1100-1106, 1989
 A:Title: The primary structure of human EGF produced by genetic engineering, studied by
 A:Reference number: A33517; MUID:89391964; PMID:2789514
 A:Accession: A33517
 A:Molecule type: protein
 R:Tsukumo, K.; Nakamura, H.; Sakamoto, S.
 Biochem. Biophys. Res. Commun. 145, 126-133, 1987
 A:Title: Purification and characterization of high molecular weight human epidermal growth
 A:Reference number: A29721; MUID:87241488; PMID:3297054
 A:Accession: A29721
 A:Molecule type: protein
 A:Residues: 829-834, 'X', 836-839, 'X', 841-845, 'X', 847-848 <TSU>
 A:Note: this is the amino-terminal sequence of a high molecular weight form of EGF, isolated
 R:Svoboda, M.; Bauhofer, A.; Schwind, P.; Bade, E.; Rasched, I.; Przybylski, M.
 Biochim. Biophys. Acta 1206, 35-41, 1994
 A:Title: Structural characterization and biological activity of recombinant human epidermal
 A:Reference number: S45282; MUID:94242778; PMID:8186248
 A:Accession: S45282
 A:Molecule type: protein
 A:Residues: 'M', 971-1023 <SVO>
 A:Note: expressed recombinant protein
 A:Accession: S45283
 A:Molecule type: protein
 A:Residues: 'MKYP', 970-1023 <SV2>
 A:Note: expressed recombinant protein
 C:Comment: Epidermal growth factor (EGF) stimulates the proliferation and differentiation of
 gastrointestinal cell proliferation.
 C:Comment: EGF is released in the pancreas, small intestine, mammary gland, and (in some
 C:Comment: The EGF precursor is found in kidney as a receptor-like membrane-bound protein
 C:Genetics:
 A:Gene: GDB:EGF
 A:Cross-references: GDB:119105; OMIM:131530
 A:Map position: 4q25-4q25
 A:Introns: 43/1; 109/3; 170/2; 246/2; 314/1; 356/3; 397/1; 438/1; 480/1; 525/3; 575/2; 6
 C:Superfamily: epidermal growth factor precursor; EGF homology; LDL receptor WYTD-contain
 C:Keywords: duplication; growth factor; tandem repeat; transmembrane protein
 F:1-22/Domain: signal sequence #status predicted <SIG>
 F:23-1207/Product: epidermal growth factor proprotein, membrane-bound form #status predi
 F:23-1032/Domain: extracellular #status predicted <EXT>
 F:43-479/Region: EGF precursor long repeat <LR1>

F:46-85/Domain: LDL receptor WYTD-containing repeat homology <YW01>
 F:86-127/Domain: LDL receptor WYTD-containing repeat homology <YW02>
 F:128-169/Domain: LDL receptor WYTD-containing repeat homology <YW03>
 F:170-211/Domain: LDL receptor WYTD-containing repeat homology <YW04>
 F:212-256/Domain: LDL receptor WYTD-containing repeat homology <YW05>
 F:257-301/Domain: LDL receptor WYTD-containing repeat homology <YW06>
 F:318-354/Domain: EGF homology <EG1>
 F:360-395/Domain: EGF homology <EG2>
 F:401-436/Domain: EGF homology <EG3>
 F:439-476/Domain: EGF homology <EG4>
 F:480-954/Region: EGF precursor long repeat <LR2>
 F:483-523/Domain: LDL receptor WYTD-containing repeat homology <YW07>
 F:524-566/Domain: LDL receptor WYTD-containing repeat homology <YW08>
 F:567-609/Domain: LDL receptor WYTD-containing repeat homology <YW09>
 F:610-653/Domain: LDL receptor WYTD-containing repeat homology <YW10>
 F:654-694/Domain: LDL receptor WYTD-containing repeat homology <YW11>
 F:695-737/Domain: LDL receptor WYTD-containing repeat homology <YW12>
 F:745-780/Domain: EGF homology <EG5>
 F:835-868/Domain: EGF homology <EG6>
 F:874-910/Domain: EGF homology <EG7>
 F:916-951/Domain: EGF homology <EG8>
 F:971-1023/Product: epidermal growth factor #status experimental <EGF>
 F:976-1012/Domain: EGF homology <EG9>
 F:1033-1057/Domain: transmembrane #status predicted <TMW>
 F:1058-1207/Domain: intracellular #status predicted <INT>
 F:318-330, 325-339, 341-354, 360-371, 367-380, 382-395, 401-412, 408-421, 423-436, 439-451, 447-461
 fide bonds: #status predicted
 F:976-990, 984-1001, 1003-1012/Disulfide bonds: #status experimental

Query Match 9.4%; Score 824.5; DB 1; Length 1207;
 Best Local Similarity 26.1%; Pred. No. 1.1e-43;
 Matches 210; Conservative 149; Mismatches 247; Indels 199; Gaps 26;

QY 12 LLLLLLLLLLALC-----GCPAPAAA-----SPLLLFANRRDRLVD 47
 Db 1 MLTLIIILLPVVSKFSFVLSAPQHSCEPTLAGNGNSTCVPAPPLIFSHGNSIFRID 60
 QY 48 AGGVKLESTIVVSGLEDAADVDFQSKGAVVWTVSEBAIKQTVLNQNGAAVQVV-1SG 106
 Db 61 TEGTNYQLVVDAGV--SVIMDFHYNEKRIIYVWDLEROLLQVFLN--GSRQERVNIEK 116
 QY 107 LVSPDGLACDWGKKLYWTDSETRIEVANINGTSRKVLFWQDLDPQRAIALDPAHGYMY 166
 Db 117 NVS--GMAINWINEVINSQOEGLITVDMKGNNSHILL-SALKYPANVAVDPERIF 173
 QY 167 WTDWGETPRIERAGMDGSTRKIIIVDSDIYWPNGITIDLEOKLYW----ADAKLSFIHRA 222
 Db 174 WSS-EVAGSLYRADLDGVGKALLET-S-EKITAVSLDVLDRFLWFIQYNREGSNLSICSC 231
 QY 223 NLDSFRQKVVEGSL---THP-----PALTISGDTLYWTDWQTSIHACNKRIG- 268
 Db 232 DYDG-----GSVHISKHPTQNLFAMSLFGDRIFVYSTWMTKIWTIANKHTGDMVRI 283
 QY 269 -----GK-----RKEILSALYSMDIQ----- 285
 Db 284 NLHSSFVPLGELKVVHFLAQPKAEDDTWEPQKCLKRGKNCSTVCGQDLQSHLCMAE 343
 QY 286 --VLSSQERQ-----PPFHT-----RCEE----- 301
 Db 344 GYALSRDRKYCEDVNECAFNNHNGCTLGCKNTPGSVYCTCPGVFLLPDGKCHQLVSCP 403
 QY 302 -----DNGGCSHLCL-LSPSEFFYTC 321
 Db 404 NVSECSHDCVLTSRGPLCFCEPGSVLBERDGKTCGSCSPDNGGCSQLCVLSPVS--WEC 461
 QY 322 ACPTGVOLQNGRCKA-GAEVILLARRTDLRLISLDTDFTDIVLQVDDIRHAIAIDY 380
 Db 462 DCFEYDLQDEKSCAASGGPQFLLFANSQDIRMHFDGTDYGLLSQ--QMGVYALDH 519
 QY 381 DPLEGYVWTDDEVAIRRAIRRAYLDGSGAQTLVNTINDPDGIADVAVARNLWTTDTGTDR 440
 Db 520 DPVENKIYFAHTALKWIERANMDCSQERLEIEGVVDYVEGLAVDMIGRRFWYTDGRKSLI 579

Qy 441 EVTRLNGTSRKILVSELDDEPRATIALHPVGMGLMYWTDGKGNPKIECANLDGQERRVLVA 500
Db 580 GRSDLNGKRSKILTYKENSIFQRGTAHVHMAKRLFWTDGIMPRIESSLQGLRULVIASS 639
Qy 501 SLGWPNGIALDLQEGKLYWGDAKTDKIEVINVDGTFKRRTLLDKLPHIFGFTLLGDFIYW 560
Db 640 DLWPSGITIDFLTDKLYWCAKQSVIEWANLDGSKRRRLTQNDVGHFFAVAVFEDYVWF 699
Qy 561 TDWQRSTSRVHVKASRDVIDQLPDLM----GLKANVAVKVTNPNCDRNGGCSHLC 616
Db 700 SDWAMPVSIRVNK-RTGKDRV--RLQGSMLKPSLSLVVHPLAKPGADPCLYQNGGCEHC 756
Qy 617 FFTPHATRCGCPIGLELLSDMKTCI 641
Db 757 KRLGTAWCSREGPMKASDKTCL 781

RESULT 15
EGMSMG
epidermal growth factor precursor - mouse
N/Alternate names: urogastone precursor
C/Species: Mus musculus (house mouse)
C/Date: 30-Nov-1980 #sequence revision 11-Aug-1983 #text_change 09-Jul-2004
R/Accession: A94272; A93304; A92118; A01387
R/Scott, J.; Urdea, M.; Quiroga, M.; Sanchez-Pescador, R.; Fong, N.; Selby, M.; Rutter, Science 221, 236-240, 1983
A/Title: Structure of a mouse submaxillary messenger RNA encoding epidermal growth factor
A/Reference number: A94272; MUID:83223630; PMID:6602382
A/Accession: A94272
A/Molecule type: mRNA
A/Residues: 1-1217 <SCO>
A/Cross-references: UNIPROT:P01132; GB:J00380; NID:g192993; PIDN:AAA37539.1; PID:g309210
R/Gray, A.; Dull, T.; Ullrich, A.
Nature 303, 722-725, 1983
A/Title: Nucleotide sequence of epidermal growth factor cDNA predicts a 128,000-molecular weight protein
A/Reference number: A93304; MUID:83219309; PMID:6304537
A/Accession: A93304
A/Molecule type: mRNA
A/Residues: 1-789, 'Y', 791-1047, 'S', 1049-1168 <GRA>
A/Cross-references: GB:J00380
A/Note: the sequence shown by these authors differs from residues 1134-1168 due to an inference of Scott et al.)
R/Savage Jr., C.R.; Inagami, T.; Cohen, S.
J. Biol. Chem. 247, 7612-7621, 1972
A/Title: The primary structure of epidermal growth factor.
A/Reference number: A92118; MUID:73048516; PMID:4636327
A/Accession: A92118
A/Molecule type: protein
A/Residues: 977-1029 <SAV>
A/Note: residues 1024-1029 are not required for full biological activity in vivo
R/Savage Jr., C.R.; Hash, J.H.; Cohen, S.
J. Biol. Chem. 248, 7669-7672, 1973
A/Title: Epidermal growth factor. Location of disulfide bonds.
A/Reference number: A92144; MUID:74025498; PMID:4750422
A/Contents: annotation; disulfide bonds
C/Comment: Epidermal growth factor (EGF) stimulates the proliferation and differentiation of gastrointestinal cell proliferation.
C/Comment: EGF is released in the pancreas, small intestine, mammary gland, and (in some species) in the kidney as a receptor-like membrane-bound protein.
C/Comment: The active growth factor from this submaxillary gland protein stimulates the proliferation of epidermal cells.
C/Superfamily: epidermal growth factor precursor; EGF homology; LDL receptor WYTD-containing repeat; duplication; growth factor; tandem repeat; transmembrane protein
F/1-28/Domain: signal sequence #status predicted <SIG>
F/29-1217/Product: epidermal growth factor precursor, membrane-bound form #status predicted <EXT>
F/29-1038/Domain: extracellular #status predicted <EXT>
F/50-485/Region: EGF precursor long repeat
F/53-92/Domain: LDL receptor WYTD-containing repeat homology <YW01>
F/93-134/Domain: LDL receptor WYTD-containing repeat homology <YW02>
F/135-176/Domain: LDL receptor WYTD-containing repeat homology <YW03>
F/177-217/Domain: LDL receptor WYTD-containing repeat homology <YW04>
F/218-262/Domain: LDL receptor WYTD-containing repeat homology <YW05>
F/263-307/Domain: LDL receptor WYTD-containing repeat homology <YW06>

F/324-360/Domain: EGF homology #status atypical <EG1>
F/366-401/Domain: EGF homology <EG2>
F/407-442/Domain: EGF homology <EG3>
F/445-482/Domain: EGF homology <EG4>
F/486-961/Region: EGF precursor long repeat
F/489-529/Domain: LDL receptor WYTD-containing repeat homology <YW07>
F/530-572/Domain: LDL receptor WYTD-containing repeat homology <YW08>
F/573-615/Domain: LDL receptor WYTD-containing repeat homology <YW09>
F/616-659/Domain: LDL receptor WYTD-containing repeat homology <YW10>
F/660-700/Domain: LDL receptor WYTD-containing repeat homology <YW11>
F/701-743/Domain: LDL receptor WYTD-containing repeat homology <YW12>
F/751-786/Domain: EGF homology <EG5>
F/842-875/Domain: EGF homology <EG6>
F/881-917/Domain: EGF homology <EG7>
F/923-958/Domain: EGF homology <EG8>
F/977-1029/Product: epidermal growth factor #status experimental <EGF>
F/982-1018/Domain: EGF homology <EG9>
F/1039-1063/Domain: transmembrane #status predicted <TMM>
F/1064-1217/Domain: intracellular #status predicted <INT>
F/347-360,366-377,373-386,388-401,407-418,414-427,429-442,445-457,453-467,469-482,751-761
tus predicted
F/982-996,990-1007,1009-1018/Disulfide bonds: #status experimental

Query Match 9.2%; Score 805; DB 1; Length 1217;
Best Local Similarity 23.9%; Pred. No. 1.8e-42;
Matches 236; Conservative 159; Mismatches 324; Indels 268; Gaps 33;

Qy 8 PPWPLLLLLLLLLALC-----GC-PAP-----AAASPLLLFANRDRVLV 46
Db 7 PTWLLAFLLVFLKISILSVTAWQTGNCQCPGLRSRSGTCAGPAPLVPFSQGSISRI 66
Qy 47 DAGVKLESTIVSGLDEAAAVDPQSKGAVYWDVSEEAIKOTYLNQGT-----AAVQNV 102
Db 67 DPDGTNHQQLVVDAGI--SADMDIHKERYIYWDVERQVLRVFLNGTGLKVCNVERK 124
Qy 103 VISGLVSPDGLACDWGKGLYWTDSNRIEVANLNGTSRKVLFWQDLDQPRATLDPAH 162
Db 125 V-----SGLAIDWIDDEVLVWQNGQVITVTDTKNSRVLL--SSLKHPNSIAVDPIE 176
Qy 163 GYMTWTWGETPRTERAGMDGSTRKIIVDSIYPNG-----LTIDLEOKLYW-----ADAK 215
Db 177 RLMPWSS--EVTGSLHRAHLKGVDTKLTET-----GGISVLTLDVLKRLFWQDSGEGS 230
Qy 216 LSFTHRANLDGSPKQVVEGSLTHPF-ALTLSDGTLVWTDWQTSIHACNKTG-----268
Db 231 HAYIHCDEYEGG-SVRLIRHQARHLSMAFFGDRIFVSVLKSKAIWANKHTGKOTVRI 289
Qy 269 -----GK-----RKEILSALYSPMDIQVLSQ-----BRQPFHT--297
Db 290 NLHPSFVTPGKLMVHPRAQPRTE--DAAKDP-DPELLKQGRPCRCGLCERDPKSHSSA 346
Qy 298 -----RCE-----301
Db 347 CAEGYTLSDRKDYCEDVNECATQNHGCTLGCENTPGSYHCTCPTGTVLLPDGKQCHELV 406
Qy 302 -----DNGGSHLCLLSPSEP-F 318
Db 407 CPGNVSKSHGCVLTSQPRICIPAGSVLGRDGTCTGCSSPDNGGCSQICL--PLRFGS 464
Qy 319 YTCACPTGVQLQDNGRTCKA-GAEVEILLARTRDLRISLDPDTPDITVLQVDDIRHAIA 377
Db 465 WECDFPGYDLQSDRKSAAASGPQPLLLFANSQDIRHMHFDGTDYK--VLLSRQGMVFA 522
Qy 378 IDYDPLEGYVWTDDEVRATRAYLDGSGAGTLVNTINDPDGTAVDVWVARNLYWTDGT 437
Db 523 LDYDPEVSKYFAQTALKIRANNDGSRERLTGEGVDTLLEGALDWIGRIYWTDSGK 582
Qy 438 DRIEVLRTNGTSRKLVSDELDPRATLHPVGMGLMYWTDGKGNPKIECANLDGQERRVL 497
Db 583 SVVGSGLSGKHRIIQERISRPRGTAVHPARLFWTDVGMSPRIESASLQSDRVLI 642
Qy 498 VNASLWPNGLALDQEGKLYWGDAKTDKIEVINVDGTFKRRTLLDKLPHIFGFTLLGDF 557

Db	643	ASSNLEPSGITIDYLTDTLYWC	DKRSVIEANLDGSKRRRLIQNDVGH	PFLAVFEDH	702	
Qy	558	IYMTDQRRS	IERVHK-----VKASRDV	IIDQLPDLMGLKAVNVAKVGTNP	604	
Db	703	LWVSDMAIP	SIVRNKRTGQNRVRLQGSMLP	SSLVVVHPL-----AK-PGADP	750	
Qy	605	CADRNGGCH	LCFFTPHATRCGCPIGLELLSDMK	TCIVPEAFVFTSRAA--IHRISLET	662	
Db	751	CLYRNGGCE	HCQESLGTARCLCREGFVKAWD	GKML-PQDYPILSGENADLSKEVTLS	809	
Qy	663	NNNDVAIPL	TGVKEASALDPDVSNHHIYWT	DVSLKTSRAFNGSSVEHVVEFGLDYPEG	722	
Db	810	NSTQAEV	PDGDTESSTL-----VAEIMV	SGMYEEDD	841	
Qy	723	MAVDMMGN	LYWADTGTNR	IEVARLDGQFRQ-----VLVWRDL	DNPRSLALDPTK	772
Db	842	CGPGCGSH	ARCVDG-ETAEQCCLKGFARD	GNLCSDIDECVLARSDCPSTSR	CINTEG	900
Qy	773	GYIYWT	EWGKPRIVRAFM	DGTCNMTL	799	
Db	901	GYVCR	SEGEYEG-----DGISC	FDI	920	

Search completed: February 17, 2005, 01:35:57 .
Job time : 69 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 17, 2005, 01:35:01 ; Search time 180 Seconds
(without alignments)
2931.666 Million cell updates/sec

Title: US-09-931-375A-2
Perfect score: 8740
Sequence: 1 MEAAPGPPWPPLLLLLLLLLL.....TERSYFHLFPFPPSPCTDSS 1615

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1376875 seqs, 326749119 residues

Total number of hits satisfying chosen parameters: 1376875

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8740	100.0	1615	10	US-09-931-375A-2
2	8736	100.0	1615	14	US-10-331-907-3
3	8736	100.0	1615	14	US-10-331-907-3
4	8736	100.0	1615	16	US-10-464-368-82
5	8724	99.8	1627	13	US-10-789-378-50
6	8715	99.7	1615	15	US-10-087-192-1410
7	8715	99.7	1615	15	US-10-374-979-3
8	8715	99.7	1615	16	US-10-182-936A-3
9	8706	99.6	1615	15	US-10-731-739-3
10	8706	99.6	1615	15	US-10-374-979-4
11	8706	99.6	1615	16	US-10-182-936A-4
12	8705	99.6	1611	15	US-10-731-739-4
13	8607	98.5	1591	14	US-10-464-368-81
					US-10-331-907-4

14	8607	98.5	1591	14	US-10-331-907-43	Sequence 43, Appl
15	8574	98.1	1639	14	US-10-331-907-29	Sequence 29, Appl
16	8568	98.0	1584	14	US-10-331-907-39	Sequence 39, Appl
17	8259.5	94.5	1614	9	US-09-887-540A-2	Sequence 2, Appl
18	8259.5	94.5	1614	14	US-10-331-907-42	Sequence 42, Appl
19	8259.5	94.5	1614	15	US-10-464-368-80	Sequence 80, Appl
20	8254.5	94.4	1614	15	US-10-464-368-75	Sequence 75, Appl
21	8254.5	94.4	1614	15	US-10-464-368-94	Sequence 94, Appl
22	8208	93.9	1586	14	US-10-464-368-74	Sequence 44, Appl
23	7882	90.2	1451	14	US-10-331-907-44	Sequence 44, Appl
24	6153.5	70.4	1613	15	US-10-331-907-25	Sequence 25, Appl
25	6097.5	69.8	1613	15	US-10-464-368-84	Sequence 84, Appl
26	4635	53.0	1494	14	US-10-464-368-83	Sequence 83, Appl
27	4635	53.0	1494	15	US-10-017-161-1612	Sequence 1612, Ap
28	3759	43.0	737	13	US-10-292-798-1286	Sequence 1286, Ap
29	2937	33.6	550	14	US-10-087-192-1407	Sequence 1407, Ap
30	2685	30.7	1576	16	US-10-331-907-7	Sequence 7, Appl
31	2685	30.7	1596	15	US-10-479-875-2	Sequence 2, Appl
32	2685	30.7	1596	15	US-10-291-172-711	Sequence 711, App
33	2685	30.7	1905	16	US-10-221-278-711	Sequence 711, App
34	2684	30.7	1637	15	US-10-480-172-6	Sequence 6, Appl
35	2684	30.7	1637	15	US-10-291-172-335	Sequence 335, App
36	2681	30.7	1852	15	US-10-221-278-335	Sequence 335, App
37	2645.5	30.3	533	14	US-10-085-198-60	Sequence 60, Appl
38	2245.5	25.7	1553	15	US-10-331-907-8	Sequence 8, Appl
39	2193	25.1	4544	15	US-10-415-188-5	Sequence 5, Appl
40	2193	25.1	4544	16	US-10-464-368-68	Sequence 68, Appl
41	2193	25.1	4563	15	US-10-479-875-6	Sequence 6, Appl
42	2173.5	24.9	4485	16	US-10-276-774-1723	Sequence 1723, Ap
43	2171.5	24.8	2641	16	US-10-741-601-332	Sequence 332, App
44	2165	24.8	4545	15	US-10-741-601-333	Sequence 333, App
45	2165	24.8	4545	15	US-10-464-368-67	Sequence 67, Appl
					US-10-464-368-71	Sequence 71, Appl

ALIGNMENTS

RESULT 1
US-09-931-375A-2
; Sequence 2, Application US/09931375A
; Publication No. US20030027151A1
; GENERAL INFORMATION:
; APPLICANT: WARMAN, Matthew L.
; APPLICANT: GONG, Yaogin
; APPLICANT: OLSEN, Bjorn R.
; APPLICANT: ROMANI, Georges
; APPLICANT: ROMAN-ROMAN, Sergio
; TITLE OF INVENTION: REGULATOR GENE AND SYSTEM USEFUL FOR THE DIAGNOSIS AND THERAPY OF
; FILE REFERENCE: 38464-0004
; CURRENT APPLICATION NUMBER: US/09/931.375A
; CURRENT FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: US 60/304,851
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/234,337
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US 60/226,119
; PRIOR FILING DATE: 2000-08-18
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 1615
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-931-375A-2

Query Match 100.0%; Score 8740; DB 10; Length 1615;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1615; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MEAAPGPPWPPLLLLLLLLLLALCCGCPAPAAAPLFFANRRDRLVLDAGGVKLESTIVVS 60
Db 1 MEAAPGPPWPPLLLLLLLLLLALCCGCPAPAAAPLFFANRRDRLVLDAGGVKLESTIVVS 60

QY 61 GLEDAADVDFQSKGAVYVTDVSEBAIKOTYLNQTAAGAVQNVVVISGLVSPDGLACDWGK 120
DB 61 GLEDAADVDFQSKGAVYVTDVSEBAIKOTYLNQTAAGAVQNVVVISGLVSPDGLACDWGK 120
QY 121 KLYWTDSETRJEVANLNGTSKRVLFWDQDOPRAIALDPAHGYMWTWDGTPRIERAG 180
DB 121 KLYWTDSETRJEVANLNGTSKRVLFWDQDOPRAIALDPAHGYMWTWDGTPRIERAG 180
QY 181 MDGSTRKIIVDSIYWPNGLTIDLEEQKLYWADAKLSFIHRANLDGSRQKVVESGLTHP 240
DB 181 MDGSTRKIIVDSIYWPNGLTIDLEEQKLYWADAKLSFIHRANLDGSRQKVVESGLTHP 240
QY 241 FALTLSGDTLYWTDQTRSIHACNKRKTKGKREIISALYSPMDIQVLQOEROPPFHTTRCE 300
DB 241 FALTLSGDTLYWTDQTRSIHACNKRKTKGKREIISALYSPMDIQVLQOEROPPFHTTRCE 300
QY 301 EDNGGCSHLCLLSPSEPFYTCAPTGVOLQDNGRTCKAGAEVLLARRDLRRLSLDTP 360
DB 301 EDNGGCSHLCLLSPSEPFYTCAPTGVOLQDNGRTCKAGAEVLLARRDLRRLSLDTP 360
QY 361 DFTDIVLQVDDIRHAIADIDPLEGYVYVWTDDEVRAIRRAYLDGSGAOTLVNTEINDPDG 420
DB 361 DFTDIVLQVDDIRHAIADIDPLEGYVYVWTDDEVRAIRRAYLDGSGAOTLVNTEINDPDG 420
QY 421 IAVDMVARNLYWTDGTDRIEVTRNGTSRKILVSEDLDEPRAIALHPVGMGLMYWTDNGE 480
DB 421 IAVDMVARNLYWTDGTDRIEVTRNGTSRKILVSEDLDEPRAIALHPVGMGLMYWTDNGE 480
QY 481 NPKIECANLDGQERRVLYNASLWGNGLALDLOEGKLYWDAKTDKIEVINVDGTRKRTL 540
DB 481 NPKIECANLDGQERRVLYNASLWGNGLALDLOEGKLYWDAKTDKIEVINVDGTRKRTL 540
QY 541 LEDKLPHPFGTLLGDFIYVTDWQRRSIRVHKVKAARDVIIDQLPDLMLGLKAVNAVAV 600
DB 541 LEDKLPHPFGTLLGDFIYVTDWQRRSIRVHKVKAARDVIIDQLPDLMLGLKAVNAVAV 600
QY 601 GTNPCADRNGGSHLCFFTPHATRCGPIGLELSDMTKICVPEAFIVFTSRAAHLRLSL 660
DB 601 GTNPCADRNGGSHLCFFTPHATRCGPIGLELSDMTKICVPEAFIVFTSRAAHLRLSL 660
QY 661 ETNNNDVAIPTGVKEASALDFVSNHHIYVTDVSLKTIISRAFMNGSSVEHVFEGLDYP 720
DB 661 ETNNNDVAIPTGVKEASALDFVSNHHIYVTDVSLKTIISRAFMNGSSVEHVFEGLDYP 720
QY 721 EGMADVMMGNLYWADTGTNRIEVARLDGQPROVLVWRDLNPRSLALDPTKGYIYWTW 780
DB 721 EGMADVMMGNLYWADTGTNRIEVARLDGQPROVLVWRDLNPRSLALDPTKGYIYWTW 780
QY 781 GKGPRIVRAFMGDTNMTLVKVRANDLTIDYADQRLYTDLDTNMTIENSNMLGQERVV 840
DB 781 GKGPRIVRAFMGDTNMTLVKVRANDLTIDYADQRLYTDLDTNMTIENSNMLGQERVV 840
QY 841 IADDLPHFPGLTQVSDYIYVTDWNLHSTIERADKTSGRNRTLIQGHLDPVMIDLVPFHSRQ 900
DB 841 IADDLPHFPGLTQVSDYIYVTDWNLHSTIERADKTSGRNRTLIQGHLDPVMIDLVPFHSRQ 900
QY 901 DGLNDCHNNGCGOLCLAIIPGHRCCGASHYTLDPSSRNCSPPTTLLFSQKSAISRM 960
DB 901 DGLNDCHNNGCGOLCLAIIPGHRCCGASHYTLDPSSRNCSPPTTLLFSQKSAISRM 960
QY 961 PDDQHSPLILPLHGLRNVAKIDYDPLDKFYIYWDGQNKIKRAKDDGTQPPVLTSLSQGG 1020
DB 961 PDDQHSPLILPLHGLRNVAKIDYDPLDKFYIYWDGQNKIKRAKDDGTQPPVLTSLSQGG 1020
QY 1021 NPDQPHDLSIDIYSRTLFTWCEATNTINVRHLSGEAMGVVLRGRDKPRAIVVNAERG 1080
DB 1021 NPDQPHDLSIDIYSRTLFTWCEATNTINVRHLSGEAMGVVLRGRDKPRAIVVNAERG 1080
QY 1081 LYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVVALVVDNTLGLFWVDADAKRIESCD 1140
DB 1081 LYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVVALVVDNTLGLFWVDADAKRIESCD 1140

QY 1141 LSGANRLTLEDANIVQPLGLTILGKHLWIDRQOQMIERVEKTTGDKXTRIQGRVAHLTG 1200
DB 1141 LSGANRLTLEDANIVQPLGLTILGKHLWIDRQOQMIERVEKTTGDKXTRIQGRVAHLTG 1200
QY 1201 IHAVEEVSLEBSAHPCARDNGGSHICIAKGDGTGPRCSVHLVLLQNLITCGSPPTCS 1260
DB 1201 IHAVEEVSLEBSAHPCARDNGGSHICIAKGDGTGPRCSVHLVLLQNLITCGSPPTCS 1260
QY 1261 PDQFACATGEIDCIPGAWRCDFPECCDDQSDDEEGCPVCSAAQFPCCARQCQVDLRLRCDGE 1320
DB 1261 PDQFACATGEIDCIPGAWRCDFPECCDDQSDDEEGCPVCSAAQFPCCARQCQVDLRLRCDGE 1320
QY 1321 ADCQDRSEADCAICLPNQPRCAGQCVLIKQOCDSPDCIDGSDDELMCEITKPPSDS 1380
DB 1321 ADCQDRSEADCAICLPNQPRCAGQCVLIKQOCDSPDCIDGSDDELMCEITKPPSDS 1380
QY 1381 PAHSSAIGPVIGIILSLFVWGGVYVFCORVVCORVAGANGPPHYYVSGTDPHVPINFTAP 1440
DB 1381 PAHSSAIGPVIGIILSLFVWGGVYVFCORVVCORVAGANGPPHYYVSGTDPHVPINFTAP 1440
QY 1441 GGSQHGPTGTIACGKSMSSVSLMGRGGVFLYDRNHVTGASSSSSSSTKATLYPPILNP 1500
DB 1441 GGSQHGPTGTIACGKSMSSVSLMGRGGVFLYDRNHVTGASSSSSSSTKATLYPPILNP 1500
QY 1501 PPSPATDPSLYNMDMFYSSNIPATARPYPYIIIRGMAPPPTTPTCSTDVCDSDYSASRWKAS 1560
DB 1501 PPSPATDPSLYNMDMFYSSNIPATARPYPYIIIRGMAPPPTTPTCSTDVCDSDYSASRWKAS 1560
QY 1561 KYIYLDLNSDSDPYPPPTPHSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615
DB 1561 KYIYLDLNSDSDPYPPPTPHSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615

RESULT 2

US-10-331-907-3

; Sequence 3, Application us/10331907

; Publication No. US20030181660A1

; GENERAL INFORMATION:

; APPLICANT: Todd, John A

; Hess, John W

; Caskey, Charles T

; Cox, Roger D

; Gerhold, David

; Hammond, Holly

; Hey, Patricia

; Kawaguchi, Yoshihiko

; Merriman, Tony R

; Metzker, Michael L

; TITLE OF INVENTION: No. US20030181660A1e1 LDL-Receptor

; NUMBER OF SEQUENCES: 455

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Nixon and Vanderhye

; STREET: 1100 No. US20030181660A1e1 Glebe Road, Eighth Floor

; CITY: Arlington

; STATE: Virginia

; COUNTRY: US

; ZIP: VA 22201-4714

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/331,907

; FILING DATE: 31-Dec-2002

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/09/402,923A

; FILING DATE: 14-Feb-2001

; APPLICATION NUMBER: PCT/GB98/01102

; FILING DATE: 15-APR-1998

; APPLICATION NUMBER: US 60/043,553

; FILING DATE: 15-APR-1997

; APPLICATION NUMBER: US 60/048,740

FILING DATE: 05-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: B. J. Sadoff
REGISTRATION NUMBER: 36,663
REFERENCE/DOCKET NUMBER: 620-81
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4091
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1615 amino acids
TYPE: amino acid
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-331-907-3

Query Match 100.0%; Score 8736; DB 14; Length 1615;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1614; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	MEAAPGPPPLLLLLLLLLLALCCCPAPAAASPLLLPANRRDRLVLDAGGVKLESTIVVS	60
DB	1	MEAAPGPPPLLLLLLLLLLALCCCPAPAAASPLLLPANRRDRLVLDAGGVKLESTIVVS	60
QY	61	GLEDAAVDFQFSKGVYWTVDVSEEAIKQTYLNQDQRAIALDPAHGYMYWTDWGETPRIERAG	120
DB	61	GLEDAAVDFQFSKGVYWTVDVSEEAIKQTYLNQDQRAIALDPAHGYMYWTDWGETPRIERAG	120
QY	121	KLWYDSETRIEVANLNGTSRKVLFWQDLDQRAIALDPAHGYMYWTDWGETPRIERAG	180
DB	121	KLWYDSETRIEVANLNGTSRKVLFWQDLDQRAIALDPAHGYMYWTDWGETPRIERAG	180
QY	181	MDGSTRKIIIVDSDIYWPNGLTIDLEOKLYWADAKLSFIHRANLDGFRKQVVEGSLTHP	240
DB	181	MDGSTRKIIIVDSDIYWPNGLTIDLEOKLYWADAKLSFIHRANLDGFRKQVVEGSLTHP	240
QY	241	FALTLSDGTLTYWTDQTRSIIHACNKRKGKRKEILSALYSPMDIQVLSQERQPFHTRCE	300
DB	241	FALTLSDGTLTYWTDQTRSIIHACNKRKGKRKEILSALYSPMDIQVLSQERQPFHTRCE	300
QY	301	EDNGGCSHLCLLSPSEPFYTCACPTGVLQNGRTCKAGAEVLLLRRTDLRRISLDTDP	360
DB	301	EDNGGCSHLCLLSPSEPFYTCACPTGVLQNGRTCKAGAEVLLLRRTDLRRISLDTDP	360
QY	361	DFTDIVLQVDDIRHAIADVDPLGYYVTTDDEVRARRAYLDGSGAQTIVNTINDPDG	420
DB	361	DFTDIVLQVDDIRHAIADVDPLGYYVTTDDEVRARRAYLDGSGAQTIVNTINDPDG	420
QY	421	IADVARNLYWTDGTDRIEVTRNGTSRKILVSEDLDEPRALHPVWGLMYWTDWGE	480
DB	421	IADVARNLYWTDGTDRIEVTRNGTSRKILVSEDLDEPRALHPVWGLMYWTDWGE	480
QY	481	NPKIECANLQGERVRLVNASLGNPGLALDLQEGKLYWDAKTDKIEVINVDGTKRRTL	540
DB	481	NPKIECANLQGERVRLVNASLGNPGLALDLQEGKLYWDAKTDKIEVINVDGTKRRTL	540
QY	541	LEDKLPHIFGFTLLGDFIYWTDQRRSIEIRVHKYKASRDVITIDQLPDLMLGLKAVNAKV	600
DB	541	LEDKLPHIFGFTLLGDFIYWTDQRRSIEIRVHKYKASRDVITIDQLPDLMLGLKAVNAKV	600
QY	601	GTNPCADNRGCSHLCEFTPHATRCGCPGLELLSDMKTCIVPEAFVLTFSRAAHRISL	660
DB	601	GTNPCADNRGCSHLCEFTPHATRCGCPGLELLSDMKTCIVPEAFVLTFSRAAHRISL	660
QY	661	ETNNNDVAIPLTGKESALDQFVSNHIIYWTVDVSLKTIISRAFWNGSSVHEVVEGLDYP	720
DB	661	ETNNNDVAIPLTGKESALDQFVSNHIIYWTVDVSLKTIISRAFWNGSSVHEVVEGLDYP	720
QY	721	EGMAVDMGNLYWADGTNRRIEVARLDGQFQVLRDLDNPRSLALDPTKGIYIYTW	780
DB	721	EGMAVDMGNLYWADGTNRRIEVARLDGQFQVLRDLDNPRSLALDPTKGIYIYTW	780
QY	781	GGKPRIVRAFMDGTNCMTLVKVRANDLTIDYADQRLYWTDLDTNMISSNMLQSERV	840

DB	781	GGKPRIVRAFMDGTNCMTLVKVRANDLTIDYADQRLYWTDLDTNMISSNMLQSERV	840
QY	841	IADLPHPPFGLTQYSDIYYTDMNLHSIERADKTSGRNRTLIQGLDFTWMDILVPHSSRQ	900
DB	841	IADLPHPPFGLTQYSDIYYTDMNLHSIERADKTSGRNRTLIQGLDFTWMDILVPHSSRQ	900
QY	901	DGLNDCMHNNGCCQLCLAIIPGHRCCGASHYTLDPSSRNCSPPTTFLFQSKSAISMI	960
DB	901	DGLNDCMHNNGCCQLCLAIIPGHRCCGASHYTLDPSSRNCSPPTTFLFQSKSAISMI	960
QY	961	PDDQHSPLILPLHLGRNVKXAIYDPLDKFYWVDGRQNIKRAKDDGTQPFVLTSLSQG	1020
DB	961	PDDQHSPLILPLHLGRNVKXAIYDPLDKFYWVDGRQNIKRAKDDGTQPFVLTSLSQG	1020
QY	1021	NPDRQPHDLSIDIYSRTILFWTCEATNTINVHRLSGEAMGVVLRGDRDKPRAIVNAERG	1080
DB	1021	NPDRQPHDLSIDIYSRTILFWTCEATNTINVHRLSGEAMGVVLRGDRDKPRAIVNAERG	1080
QY	1081	LYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVVALVVDNTLGLFWVDADLKRIESCD	1140
DB	1081	LYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVVALVVDNTLGLFWVDADLKRIESCD	1140
QY	1141	LSGANRLTLEDANIVQPLGLTILGKLYWIDRQOQMIERVEKTTGDKRTRIQGRVAHLTG	1200
DB	1141	LSGANRLTLEDANIVQPLGLTILGKLYWIDRQOQMIERVEKTTGDKRTRIQGRVAHLTG	1200
QY	1201	IHAVEEVSLEBSAHPCARDNGGCSHICIAKGGTTPRCPCPVHLVLQNLITCGEPPTCS	1260
DB	1201	IHAVEEVSLEBSAHPCARDNGGCSHICIAKGGTTPRCPCPVHLVLQNLITCGEPPTCS	1260
QY	1261	PDQFACATGETDCIPGAWRCDFPECDQSDDEECPCVCSAAQPCARQCVDLRLRCQGE	1320
DB	1261	PDQFACATGETDCIPGAWRCDFPECDQSDDEECPCVCSAAQPCARQCVDLRLRCQGE	1320
QY	1321	ADCDRSDRDEADCAICLPNQFRCSAGQCVLIKQCCDSFPDCIDGSDDELWCBITPPSDDS	1380
DB	1321	ADCDRSDRDEADCAICLPNQFRCSAGQCVLIKQCCDSFPDCIDGSDDELWCBITPPSDDS	1380
QY	1381	PAHSSAIGPVIGIILSLFVNGVYFVQVVCQRYAGANGPPHYVSGTHVPLNFIAP	1440
DB	1381	PAHSSAIGPVIGIILSLFVNGVYFVQVVCQRYAGANGPPHYVSGTHVPLNFIAP	1440
QY	1441	GGSHQGPFTGLACSKMSSVSLMGGRGVPLYDRNVHTGASSSSSSSTKATLPPILNP	1500
DB	1441	GGSHQGPFTGLACSKMSSVSLMGGRGVPLYDRNVHTGASSSSSSSTKATLPPILNP	1500
QY	1501	PPSPATDPSLYNMDMFYSSNIPATAPYRPIYIRGMAPPPTTTCSTDVCDSDYSASRWKAS	1560
DB	1501	PPSPATDPSLYNMDMFYSSNIPATAPYRPIYIRGMAPPPTTTCSTDVCDSDYSASRWKAS	1560
QY	1561	KYILDLNSDSDPPPTPHSQYLSAEDSCPPSPATERSYPHLPPPPSPCTDSS	1615
DB	1561	KYILDLNSDSDPPPTPHSQYLSAEDSCPPSPATERSYPHLPPPPSPCTDSS	1615

RESULT 3
US-10-464-368-82
; Sequence 82, Application US/10464368
; Publication No. US20040023356A1
; GENERAL INFORMATION:
; APPLICANT: Krumlauf, Robb
; APPLICANT: Ellies, Debra
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR REGULATING BONE DEPOSITION
; FILE REFERENCE: 40716-IP-017
; CURRENT APPLICATION NUMBER: US/10/464,368
; CURRENT FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: 60/388,970
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 82
; LENGTH: 1615

; TYPE: PRT									
; ORGANISM: HOMO SAPIENS									
US-10-464-368-82									
Query Match 100.0%; Score 8736; DB 15; Length 1615;									
Best Local Similarity 99.9%; Pred. No. 0;									
Matches 1614; Conservative 0; Mismatches 1; Indels 0; Gaps 0;									
QY	1	MEAAAPGPPWPLLLLLLLLLLALCGCPAPAAASPLLLFANRRDRLVLDAGGVKLESTIVVS	60						
DB	1	MEAAAPGPPWPLLLLLLLLLLALCGCPAPAAASPLLLFANRRDRLVLDAGGVKLESTIVVS	60						
QY	61	GLEDAAAVDQFSGKAVYWDVSEBAIKOTYLNQTAQVQNVVVISGLVSPDGLACDVGK	120						
DB	61	GLEDAAAVDQFSGKAVYWDVSEBAIKOTYLNQTAQVQNVVVISGLVSPDGLACDVGK	120						
QY	121	KLYWTDSETNRIEVANLNGTSKVLFWQDLPRAIALDPAGHYMYWTDWGETPRIERAG	180						
DB	121	KLYWTDSETNRIEVANLNGTSKVLFWQDLPRAIALDPAGHYMYWTDWGETPRIERAG	180						
QY	181	MDGSTRKIIVSDIYWPNGLTIDLEOKLYWADAKLSFTHRANLDGSPROKVVESGLTHP	240						
DB	181	MDGSTRKIIVSDIYWPNGLTIDLEOKLYWADAKLSFTHRANLDGSPROKVVESGLTHP	240						
QY	241	FALTLSGDTLYWTDWQTSIIHACNKRGTGKRKEIISALYSMDIOVLQSOERQPPHTRCE	300						
DB	241	FALTLSGDTLYWTDWQTSIIHACNKRGTGKRKEIISALYSMDIOVLQSOERQPPHTRCE	300						
QY	301	EDNGCCHLCLLSPSEPFYTCAPTGVQIQDNGRTCKAGABEVLLARRTDLRLSLDTP	360						
DB	301	EDNGCCHLCLLSPSEPFYTCAPTGVQIQDNGRTCKAGABEVLLARRTDLRLSLDTP	360						
QY	361	DFTDVLQVDDIRHAIADYDPLEGYVYWDDEVRAIRRAYLDGSGAOTLVNTEINDPDG	420						
DB	361	DFTDVLQVDDIRHAIADYDPLEGYVYWDDEVRAIRRAYLDGSGAOTLVNTEINDPDG	420						
QY	421	IADVAAANLYWTDGTDRIEVTRNGTSRKILVSEDLDEPRALHPVGMVMTWDMGE	480						
DB	421	IADVAAANLYWTDGTDRIEVTRNGTSRKILVSEDLDEPRALHPVGMVMTWDMGE	480						
QY	481	NPKIECANLQGERVLVNASLGNPNGLALDQEGKLYWDAKTDKIEVINVDGTKRTL	540						
DB	481	NPKIECANLQGERVLVNASLGNPNGLALDQEGKLYWDAKTDKIEVINVDGTKRTL	540						
QY	541	LEDKLPHPFGTLGDFIYWDWQRSIERVHKYKASRDVLIIDQLPDLMLKAVNAVAV	600						
DB	541	LEDKLPHPFGTLGDFIYWDWQRSIERVHKYKASRDVLIIDQLPDLMLKAVNAVAV	600						
QY	601	GTNPCADRNGGCSHLCPFTPHATRCGCPGLELLSDMKTCIVPBAFLVFTSRAAIHRISL	660						
DB	601	GTNPCADRNGGCSHLCPFTPHATRCGCPGLELLSDMKTCIVPBAFLVFTSRAAIHRISL	660						
QY	661	ETNNNDVAIPUTGVKEASALDPDVSNHHIYWDVSLKTIISRAFMNGSSVEHVFFGLDYP	720						
DB	661	ETNNNDVAIPUTGVKEASALDPDVSNHHIYWDVSLKTIISRAFMNGSSVEHVFFGLDYP	720						
QY	721	EGMAVDWNGKLYWADTGTNRIEVARLDGQROVLVWRDLNPRSLALDPTKGYIYTW	780						
DB	721	EGMAVDWNGKLYWADTGTNRIEVARLDGQROVLVWRDLNPRSLALDPTKGYIYTW	780						
QY	781	CGKPRIVRAFMGDMTCMTLVKVRANDLTIDYADQRLYWDLTDTNMISSNMLGQERVV	840						
DB	781	CGKPRIVRAFMGDMTCMTLVKVRANDLTIDYADQRLYWDLTDTNMISSNMLGQERVV	840						
QY	841	IADLPHFPGLTQYSDIYTWDMNLHSTERADKTSGRNRTLIQGHDFWMDILVFHSSRQ	900						
DB	841	IADLPHFPGLTQYSDIYTWDMNLHSTERADKTSGRNRTLIQGHDFWMDILVFHSSRQ	900						
QY	901	DGLNDCHNNGCCQOLCLAIPEGHRCGCASHYTLDPGSRNCSPTTFLFSQKSAISRM	960						
DB	901	DGLNDCHNNGCCQOLCLAIPEGHRCGCASHYTLDPGSRNCSPTTFLFSQKSAISRM	960						
QY	961	PDDQHSPLILPLHGLRNVAIDYDPLDKFYWDGRQNIKRAKDDGTPPVLTSLSQGG	1020						

RESULT 4
US-10-789-378-50
; Sequence 50, Application US/10789378
; Publication No. US20050003390A1
; GENERAL INFORMATION:
; APPLICANT: Azenovich, Sergey
; APPLICANT: Stull, Robert
; APPLICANT: Gelman, Marina
; APPLICANT: Chui, Kitty
; APPLICANT: Ng, Dean
; TITLE OF INVENTION: DIAGNOSTIC METHODS FOR CANCER DETECTION
; FILE REFERENCE: 5189-2
; CURRENT APPLICATION NUMBER: US/10789,378
; CURRENT FILING DATE: 2004-02-26
; PRIOR APPLICATION NUMBER: 10/441925
; PRIOR FILING DATE: 2003-05-19
; PRIOR APPLICATION NUMBER: 60/381619
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/450886
; PRIOR FILING DATE: 2003-02-26
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 50
; LENGTH: 1615
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-789-378-50
Query Match 100.0%; Score 8736; DB 16; Length 1615;

Best Local Similarity 99.9%; Pred. No. 0;		Matches 1614; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
Qy	1	MEAAPGPPWLLLLLLLALCCCPAPAAAPLFFANRRDVLVDAGGVKLESTIVS	60
Db	1	MEAAPGPPWLLLLLLLALCCCPAPAAAPLFFANRRDVLVDAGGVKLESTIVS	60
Qy	61	GLEDAAAVDFQSKGAVTVDVSEAIKQTYLNQGAOVNVIISGLVSPDGLACDWGK	120
Db	61	GLEDAAAVDFQSKGAVTVDVSEAIKQTYLNQGAOVNVIISGLVSPDGLACDWGK	120
Qy	121	KLYTDSGTNRIVANLNGTSRKVLFWQDLPQRAIALDPAGHYMYTWDGTPRIERAG	180
Db	121	KLYTDSGTNRIVANLNGTSRKVLFWQDLPQRAIALDPAGHYMYTWDGTPRIERAG	180
Qy	181	MDGSTRIIVDSDIYWPNGLTIDLEBOKLYWADAKLSFIHEANLDGSPROKVGSLTTP	240
Db	181	MDGSTRIIVDSDIYWPNGLTIDLEBOKLYWADAKLSFIHEANLDGSPROKVGSLTTP	240
Qy	241	FALTLSGDTLYTWDQTRSIIHACNKRITGGKKEILSALYSPMDIQVLSQERQPFHTRCE	300
Db	241	FALTLSGDTLYTWDQTRSIIHACNKRITGGKKEILSALYSPMDIQVLSQERQPFHTRCE	300
Qy	301	EDNGGCSHLCLLSPSEPFYTCACPTGVLQDNGRTCKAGAEVLLARRTDLRRISLDT	360
Db	301	EDNGGCSHLCLLSPSEPFYTCACPTGVLQDNGRTCKAGAEVLLARRTDLRRISLDT	360
Qy	361	DFTDILVQDDIRHAIADIDPLRGVYVTTDDEVRARRAYLVDGSGAQTLVNTINPDG	420
Db	361	DFTDILVQDDIRHAIADIDPLRGVYVTTDDEVRARRAYLVDGSGAQTLVNTINPDG	420
Qy	421	IADVARNLYTDTGTDRIEVTRNGTSRKILVSIEDLDEPRALHPVWGLMYTWDGGE	480
Db	421	IADVARNLYTDTGTDRIEVTRNGTSRKILVSIEDLDEPRALHPVWGLMYTWDGGE	480
Qy	481	NPKIECANLDQERRVLVNASLGNPLGLDQEGKLYWGDATDKIEVINVDGTKERTL	540
Db	481	NPKIECANLDQERRVLVNASLGNPLGLDQEGKLYWGDATDKIEVINVDGTKERTL	540
Qy	541	LEDKLPHEFGLTLDGFTYTDWQRRSIEVHVKASRDVLIIDQLPDLMLKANVNAKV	600
Db	541	LEDKLPHEFGLTLDGFTYTDWQRRSIEVHVKASRDVLIIDQLPDLMLKANVNAKV	600
Qy	601	GTNPACDRNGGCSHLCEFTTHATRCGCPIGLELLSDMKTCIVPEAFVFTSRAAIIHISL	660
Db	601	GTNPACDRNGGCSHLCEFTTHATRCGCPIGLELLSDMKTCIVPEAFVFTSRAAIIHISL	660
Qy	661	ETNNVDVAIPLTGKESALDQFVSNHHIYWTVDVSLKTIISRAFNMGSSVHVVEFGLDYP	720
Db	661	ETNNVDVAIPLTGKESALDQFVSNHHIYWTVDVSLKTIISRAFNMGSSVHVVEFGLDYP	720
Qy	721	EGMAVDWGNLYWADTGTNRIEVRLDQFQVLRDLNPSRLALDPTKGYIYVTEW	780
Db	721	EGMAVDWGNLYWADTGTNRIEVRLDQFQVLRDLNPSRLALDPTKGYIYVTEW	780
Qy	781	GGKPRIVRAFMDGTNCTLVKVRANDLTIDYADQRLYMTDLTNMIESNNMLGQSRVV	840
Db	781	GGKPRIVRAFMDGTNCTLVKVRANDLTIDYADQRLYMTDLTNMIESNNMLGQSRVV	840
Qy	841	IADLPHPPFGLTQVSDIYIYTDWNLHIERADKTSGRNRTLIOGHLDVMDILVFHSSRQ	900
Db	841	IADLPHPPFGLTQVSDIYIYTDWNLHIERADKTSGRNRTLIOGHLDVMDILVFHSSRQ	900
Qy	901	DGLNDCMHNNGCCQLCIAIPGHRCCASHYTLDPSSRNCSPPTTLLFSQKSAISRM	960
Db	901	DGLNDCMHNNGCCQLCIAIPGHRCCASHYTLDPSSRNCSPPTTLLFSQKSAISRM	960
Qy	961	PDDQHSPLILPLHGLNRVAIDYDPLDKFYWDGQNKIKRAKDDGTQPPVLTSLSQGG	1020
Db	961	PDDQHSPLILPLHGLNRVAIDYDPLDKFYWDGQNKIKRAKDDGTQPPVLTSLSQGG	1020
Qy	1021	NPDQPHDLSDIYSRTLFWTCEATNTINVHRLSGEAMGVVLRGDRDKPRAIVVNAERGY	1080

Db	1021	NPDQPHDLSDIYSRTLFWTCEATNTINVHRLSGEAMGVVLRGDRDKPRAIVVNAERGY	1080
Qy	1081	LYFTNMODRAAKIERAALDGTREVLFTTGLIRVALVVDNTLGLFWVDADLKRIEBCD	1140
Db	1081	LYFTNMODRAAKIERAALDGTREVLFTTGLIRVALVVDNTLGLFWVDADLKRIEBCD	1140
Qy	1141	LSGANRLTLEDANIVQPLGLTILGKHYLIDRQOQMIERVEKTTGDKRTRIQQRVAHLTG	1200
Db	1141	LSGANRLTLEDANIVQPLGLTILGKHYLIDRQOQMIERVEKTTGDKRTRIQQRVAHLTG	1200
Qy	1201	IHAVEEVSLEBFSAPCARDNGGCSHICIAKGDGTPRCPCPVHLVLLQNLLTCGEPPTCS	1260
Db	1201	IHAVEEVSLEBFSAPCARDNGGCSHICIAKGDGTPRCPCPVHLVLLQNLLTCGEPPTCS	1260
Qy	1261	PQFACATGETDCIIPGAWRCDFPECDQSDDEECPCVCSAAQFPCARGCCVDLRLRCDGE	1320
Db	1261	PQFACATGETDCIIPGAWRCDFPECDQSDDEECPCVCSAAQFPCARGCCVDLRLRCDGE	1320
Qy	1321	ADCDRDEADCDACILPNOFRFCASGQCVLIKQCCDSFPDCIDGSDDELMCEITKPPSDDS	1380
Db	1321	ADCDRDEADCDACILPNOFRFCASGQCVLIKQCCDSFPDCIDGSDDELMCEITKPPSDDS	1380
Qy	1381	PAHSSAIGPVIGIILSLFVWGCVYFVQVQVQRYAGANGPPHYVSGTTPHVLNFIAP	1440
Db	1381	PAHSSAIGPVIGIILSLFVWGCVYFVQVQVQRYAGANGPPHYVSGTTPHVLNFIAP	1440
Qy	1441	GGSHGPGFTGIACGKSNMSSVSLMGGRGVPLYDRNHVGTGASSSSSTKATLYPPIINP	1500
Db	1441	GGSHGPGFTGIACGKSNMSSVSLMGGRGVPLYDRNHVGTGASSSSSTKATLYPPIINP	1500
Qy	1501	PPSPATDPSLYNMDMFYSSNIPATAPYRPIYIRGMAPPTTTCSTDVCDSDYSASRWKAS	1560
Db	1501	PPSPATDPSLYNMDMFYSSNIPATAPYRPIYIRGMAPPTTTCSTDVCDSDYSASRWKAS	1560
Qy	1561	KYILDNSDSDPYPPPTPHSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS	1615
Db	1561	KYILDNSDSDPYPPPTPHSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS	1615
RESULT 5			
US-10-087-192-1410			
; Sequence 1410, Application US/10087192			
; Publication No. US20020182586A1			
; GENERAL INFORMATION:			
; APPLICANT: Engelhard, Eric K.			
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR			
; FILE REFERENCE: 529452000122			
; CURRENT APPLICATION NUMBER: US/10/087,192			
; CURRENT FILING DATE: 2002-03-01			
; PRIOR APPLICATION NUMBER: US 09/747,377			
; PRIOR FILING DATE: 2000-12-22			
; PRIOR APPLICATION NUMBER: US 09/798,586			
; PRIOR FILING DATE: 2001-03-02			
; NUMBER OF SEQ ID NOS: 2059			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 1410			
; LENGTH: 1627			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-10-087-192-1410			
Query Match 99.8%; Score 8724; DB 13; Length 1627;			
Best Local Similarity 99.3%; Pred. No. 0;			
Matches 1615; Conservative 0; Mismatches 0; Indels 12; Gaps 1;			
Qy	1	MEAAPGPPWLLLLLLLALCCCPAPAAAPLFFANRRDVLVDAGGVKLESTIVS	60
Db	1	MEAAPGPPWLLLLLLLALCCCPAPAAAPLFFANRRDVLVDAGGVKLESTIVS	60
Qy	61	GLEDAAAVDFQSKGAVYVTDVSEAIKQTYLNQGAOVNVIISGLVSPDGLACDWGK	120

Db 61 GLEDAAVDFQFSGKGVYVTDVSEAIKQTYLNTQGAQVQNVVLSGLVSPDGLACDWGK 120
Qy 121 KLYWTDSETRLEVANLNGTSKRVFWQDLDPRAIALDPAGHYWTDWGTPTRIERAG 180
Db 121 KLYWTDSETRLEVANLNGTSKRVFWQDLDPRAIALDPAGHYWTDWGTPTRIERAG 180
Qy 181 MDGSTRKIIIVSDIYWPNGLTLDLEOKLYWADAKLSPIHRLANLDSGSRQKVEGSLTHP 240
Db 181 MDGSTRKIIIVSDIYWPNGLTLDLEOKLYWADAKLSPIHRLANLDSGSRQKVEGSLTHP 240
Qy 241 FALTLSGDTLYWTDQTRSIIHACNRTGGRKEILSALYSFMDIOVLQOEROPFFHTRCE 300
Db 241 FALTLSGDTLYWTDQTRSIIHACNRTGGRKEILSALYSFMDIOVLQOEROPFFHTRCE 300
Qy 301 EONGGCSHLLSPSEPFYTCAPTGVQLQDNGRTCKAGAEVLLARLTDLRRISLDTPT 360
Db 301 EONGGCSHLLSPSEPFYTCAPTGVQLQDNGRTCKAGAEVLLARLTDLRRISLDTPT 360
Qy 361 DFTDIVLQVDDIRHAIADIDYDPLEGVVYVTDDEVRAIRAYLDGSGAOTLVNTEINDPDG 420
Db 361 DFTDIVLQVDDIRHAIADIDYDPLEGVVYVTDDEVRAIRAYLDGSGAOTLVNTEINDPDG 420
Qy 421 IAVDVARNLYWTDGTDRIEVTRNGTSRKILVSEDLDEPRAIALHPVMGLMYTWDGE 480
Db 421 IAVDVARNLYWTDGTDRIEVTRNGTSRKILVSEDLDEPRAIALHPVMGLMYTWDGE 480
Qy 481 NPKIECANLDGQRRVLYNASLGNGLALDLOEGLYWGDAKTDKIEVINVDGTRKRTL 540
Db 481 NPKIECANLDGQRRVLYNASLGNGLALDLOEGLYWGDAKTDKIEVINVDGTRKRTL 540
Qy 541 LEDKLPHPFGFTLLGDFYVTDWQRRSIEVHVKKASRDVVIDQLPDLMLKAVNAVAVV 600
Db 541 LEDKLPHPFGFTLLGDFYVTDWQRRSIEVHVKKASRDVVIDQLPDLMLKAVNAVAVV 600
Qy 601 GTNPCADRNGGCSHLCFFTPHATRCGPIGLELLSDMKTCIVPEAFVFTSRAAIIHRLS 660
Db 601 GTNPCADRNGGCSHLCFFTPHATRCGPIGLELLSDMKTCIVPEAFVFTSRAAIIHRLS 660
Qy 661 ETNNNDVAIPLTGVKEASALDPDVSNHHIYWDVSLKTIISRAFPMGSSVHVHVEFGLDYP 720
Db 661 ETNNNDVAIPLTGVKEASALDPDVSNHHIYWDVSLKTIISRAFPMGSSVHVHVEFGLDYP 720
Qy 721 EGMAVDMGKNLYWADTGTNRLEVARLDGQFQVLRVWRDLNPRSLALDPTKGYIYTWTEW 780
Db 721 EGMAVDMGKNLYWADTGTNRLEVARLDGQFQVLRVWRDLNPRSLALDPTKGYIYTWTEW 780
Qy 781 GSKPRIVAFMDGTNCMTLVDKVGSRANDLTIDYADQRLYWTDLPTNMIESNNMLGQERVV 840
Db 781 GSKPRIVAFMDGTNCMTLVDKVGSRANDLTIDYADQRLYWTDLPTNMIESNNMLGQERVV 840
Qy 841 IADDLPHFPGLTOYSDIYVTDWNLHSTIERADKTSGRNRTLIQGLHDFVMDILVPHSSRQ 900
Db 841 IADDLPHFPGLTOYSDIYVTDWNLHSTIERADKTSGRNRTLIQGLHDFVMDILVPHSSRQ 900
Qy 901 DGLNDCMHNGCGQLCLAIPEGHRCGCASHYTLDPSSRNCSPPTFFLLFSQSAISRMI 960
Db 901 DGLNDCMHNGCGQLCLAIPEGHRCGCASHYTLDPSSRNCSPPTFFLLFSQSAISRMI 960
Qy 961 PDDQHSPLILPLHGLRNVKAIDYDPLDKFYVWDGRQNIKRAKDDGT----- 1008
Db 961 PDDQHSPLILPLHGLRNVKAIDYDPLDKFYVWDGRQNIKRAKDDGTQAVLRSLGFVVL 1020
Qy 1009 QPFLVTLSSQGNPDPRPHDLSIDYSTRFTWTCENTNTINVRHLSGAMGVLRGRDCK 1068
Db 1021 QPFLVTLSSQGNPDPRPHDLSIDYSTRFTWTCENTNTINVRHLSGAMGVLRGRDCK 1080
Qy 1069 PRAIVVNAERGVLFTNNQDRAAKIERAALDGTREVLFTTGLIRPVAVVDNTLGLKFW 1128
Db 1081 PRAIVVNAERGVLFTNNQDRAAKIERAALDGTREVLFTTGLIRPVAVVDNTLGLKFW 1140
Qy 1129 VDADLKRIESCDLSANRLTLEDANI VQPLGTLTLGKHLWIDRQQQMIERVEKTTGDGR 1188
Db 1141 VDADLKRIESCDLSANRLTLEDANI VQPLGTLTLGKHLWIDRQQQMIERVEKTTGDGR 1200

Qy 1189 TRIQGRVAHLGTGIIHAEVVSLEEFSAHPCARDNGGCSHICIAKGDGTPRCPCVHLVLLQ 1248
Db 1201 TRIQGRVAHLGTGIIHAEVVSLEEFSAHPCARDNGGCSHICIAKGDGTPRCPCVHLVLLQ 1260
Qy 1249 NLLTCGEPTCSPDOFACATGEIDICIPGAWCDGPECCDDQSDDEGCPVCSAAQPPCARG 1308
Db 1261 NLLTCGEPTCSPDOFACATGEIDICIPGAWCDGPECCDDQSDDEGCPVCSAAQPPCARG 1320
Qy 1309 QCVDLRLRCDEADQCQDRSDEADCDALCLPQFRASCQCQVLIKQCCDSFPDCIDGSDDEL 1368
Db 1321 QCVDLRLRCDEADQCQDRSDEADCDALCLPQFRASCQCQVLIKQCCDSFPDCIDGSDDEL 1380
Qy 1369 MCEITKPSDDSPAHSSAIGVIGIILSLFVWGVIYVQCRVQCRVAGAGPPPHYVS 1428
Db 1381 MCEITKPSDDSPAHSSAIGVIGIILSLFVWGVIYVQCRVQCRVAGAGPPPHYVS 1440
Qy 1429 GTPHVPLNFIAPGSGQHPFTGIACGKSMSSSVSLMGRGGVPLYDRNHVGTGASSSSSS 1488
Db 1441 GTPHVPLNFIAPGSGQHPFTGIACGKSMSSSVSLMGRGGVPLYDRNHVGTGASSSSSS 1500
Qy 1489 TKATLYPILNPPSPATDPSLYNMDMFYSSNIPATARPYPYIIRGMATPTTTCSTDVC 1548
Db 1501 TKATLYPILNPPSPATDPSLYNMDMFYSSNIPATARPYPYIIRGMATPTTTCSTDVC 1560
Qy 1549 DSDYSASRWKASKYLYDLNDSDDYPPTPHSQYLSAEDSCPPSPATERSYFHLFPPPP 1608
Db 1561 DSDYSASRWKASKYLYDLNDSDDYPPTPHSQYLSAEDSCPPSPATERSYFHLFPPPP 1620
Qy 1609 SPCTDSS 1615
Db 1621 SPCTDSS 1627

RESULT 6

US-10-374-979-3
; Sequence 3, Application US/10374979
; Publication No. US20030219793A1
; GENERAL INFORMATION:
; APPLICANT: John P. Carulli et al.
; TITLE OF INVENTION: THE HIGH BONE MASS GENE OF 11q13.3
; FILE REFERENCE: 032796-021
; CURRENT APPLICATION NUMBER: US/10/374,979
; CURRENT FILING DATE: 2003-03-04
; PRIOR APPLICATION NUMBER: US 09/544,398
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 09/543,771
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 09/229,319
; PRIOR FILING DATE: 1999-01-13
; PRIOR APPLICATION NUMBER: US 60/071,449
; PRIOR FILING DATE: 1998-01-13
; PRIOR APPLICATION NUMBER: US 60/105,511
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 109
; SEQ ID NO 3
; LENGTH: 1615
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-374-979-3

Query Match 99.7%; Score 8715; DB 15; Length 1615;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 1610; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MEAAPPGPWPLLLLLLLALCGCPAPAAASPLLLFANRRDRLVDAGGVKLESTIVVS 60

Db 1 MEAAPPGPWPLLLLLLLALCGCPAPAAASPLLLFANRRDRLVDAGGVKLESTIVVS 60

Qy 61 GLEDAAVDFQFSGKGVYVTDVSEAIKQTYLNTQGAQVQNVVLSGLVSPDGLACDWGK 120

Db 61 GLEDAAVDFQFSGKGVYVTDVSEAIKQTYLNTQGAQVQNVVLSGLVSPDGLACDWGK 120

```
Qy 121 KLYWTDSETRNRIEVANLNGTSRKVLFWQDLDPQRAIALDPAHGYMYWTWGTETPRIERAG 180
Db 121 KLYWTDSETRNRIEVANLNGTSRKVLFWQDLDPQRAIALDPAHGYMYWTWGTETPRIERAG 180
Qy 181 MDGSTRKLIIVDSDIYWPNGLTIDLEQKLYWADAKLSFIHRANLDGSPRQKRVGSLTHP 240
Db 181 MDGSTRKLIIVDSDIYWPNGLTIDLEQKLYWADAKLSFIHRANLDGSPRQKRVGSLTHP 240
Qy 241 FALTSLGDTLWTDWQTRSIIHACNKRITGKKEIILSALYSPMDIQVLSOERQPFPHTRCE 300
Db 241 FALTSLGDTLWTDWQTRSIIHACNKRITGKKEIILSALYSPMDIQVLSOERQPFPHTRCE 300
Qy 301 EDNGGCSHCLLSPSEPPYTCACPTGVLQDNGRTCKAGAEVLLARTRDLRRLSLDTP 360
Db 301 EDNGGCSHCLLSPSEPPYTCACPTGVLQDNGRTCKAGAEVLLARTRDLRRLSLDTP 360
Qy 361 DPTDILVQVDDIRHAIADIDYDPLGYYVWTDDEVRATRRAYLDGSGAQTLVNTINPDG 420
Db 361 DPTDILVQVDDIRHAIADIDYDPLGYYVWTDDEVRATRRAYLDGSGAQTLVNTINPDG 420
Qy 421 IAVDWARNLYWTDGTDRIEVTRNLGTSRKILVSEDLDEPRALHPVGLMTWTWGE 480
Db 421 IAVDWARNLYWTDGTDRIEVTRNLGTSRKILVSEDLDEPRALHPVGLMTWTWGE 480
Qy 481 NPKIECANLDQERRVLVNASLGNPLALDLQEGKLYMGDAKTDKIEVINVDGTRKRTL 540
Db 481 NPKIECANLDQERRVLVNASLGNPLALDLQEGKLYMGDAKTDKIEVINVDGTRKRTL 540
Qy 541 LEDKLPHPFGTLTGDFIYWTWQRRSIEVHVKKASRDVLIIDQLPDLMLGLKAVNAKV 600
Db 541 LEDKLPHPFGTLTGDFIYWTWQRRSIEVHVKKASRDVLIIDQLPDLMLGLKAVNAKV 600
Qy 601 GTNPCADRNGCSHLCFFTHATCGCPIGLELSDMKTCIVPEAFVFTSRAAHRISL 660
Db 601 GTNPCADRNGCSHLCFFTHATCGCPIGLELSDMKTCIVPEAFVFTSRAAHRISL 660
Qy 661 ETNNNDVAIPLTGVKEASALDPDVSNNHIYWTVSLKTTISRAFNWGSSEVHVFGLDYP 720
Db 661 ETNNNDVAIPLTGVKEASALDPDVSNNHIYWTVSLKTTISRAFNWGSSEVHVFGLDYP 720
Qy 721 EGMADVWNGKLYWADTGTNRIEVARLDGQROVLVWRDLNPRSLALDPTKGYIYTW 780
Db 721 EGMADVWNGKLYWADTGTNRIEVARLDGQROVLVWRDLNPRSLALDPTKGYIYTW 780
Qy 781 GKGKRIIVRAFMDGNTCTVLDKVRANDLTIDYADQRLYWTDLTNMIESNNMLGQERV 840
Db 781 GKGKRIIVRAFMDGNTCTVLDKVRANDLTIDYADQRLYWTDLTNMIESNNMLGQERV 840
Qy 841 IADLPHFPGLTOYSYDIYWTDMNLHSIERADKTSGRNRTLIOGHLDVMDILVFHSSRQ 900
Db 841 IADLPHFPGLTOYSYDIYWTDMNLHSIERADKTSGRNRTLIOGHLDVMDILVFHSSRQ 900
Qy 901 DGLNDCHNNGCQQLCLAIPEGHRCGCASHYTLDPSSRNCSPPTTFLPSQKSAISRM 960
Db 901 DGLNDCHNNGCQQLCLAIPEGHRCGCASHYTLDPSSRNCSPPTTFLPSQKSAISRM 960
Qy 961 PDDQHSPLILPLAGLRNVAKIDYDLDKFTIYWDGRQNIKRAKDDGTOPVLTSLSQG 1020
Db 961 PDDQHSPLILPLAGLRNVAKIDYDLDKFTIYWDGRQNIKRAKDDGTOPVLTSLSQG 1020
Qy 1021 NPDQPHDLSIDIYSRTLFWTCEATNTINVHRLSGEAMGVVLRGDRDKPRAIVNARGY 1080
Db 1021 NPDQPHDLSIDIYSRTLFWTCEATNTINVHRLSGEAMGVVLRGDRDKPRAIVNARGY 1080
Qy 1081 LYFTNMQDRAAKIBRAALDGTREVLFTTGLIRPVALVVDNTLGLFWADLKRIESCD 1140
Db 1081 LYFTNMQDRAAKIBRAALDGTREVLFTTGLIRPVALVVDNTLGLFWADLKRIESCD 1140
Qy 1141 LSGANRLTLEDANIVQPLGTLCKHLYWIDRQOQMIERVEKTTGDKRTRIQGRVAHLT 1200
Db 1141 LSGANRLTLEDANIVQPLGTLCKHLYWIDRQOQMIERVEKTTGDKRTRIQGRVAHLT 1200
Qy 1201 IHAVEEVSLEBFAHPCARDNGGCSHICIAKGDGTPRCSCPVHLVLLQNLLTCGEPPTCS 1260
```

```
Db 1201 IHAVEEVSLEBFAHPCARDNGGCSHICIAKGDGTPRCSCPVHLVLLQNLLTCGEPPTCS 1260
Qy 1261 PDQFACATGEIDCTPGAWRCDFPECCDDQSDSEBECPCVCSAAQFFCARGQCQVDLRLRCDGE 1320
Db 1261 PDQFACATGEIDCTPGAWRCDFPECCDDQSDSEBECPCVCSAAQFFCARGQCQVDLRLRCDGE 1320
Qy 1321 ADCQDRDEADCDAILPNQPRCASGOCVLLIKQCCDSPDCIDGSDLMCEITKPPSDDS 1380
Db 1321 ADCQDRDEADCDAILPNQPRCASGOCVLLIKQCCDSPDCIDGSDLMCEITKPPSDDS 1380
Qy 1381 PAHSSAIGPVIGIILSLFVMGGVYFVQORVVQRYAGANGPPHYYVSGTTPHVPPLFIAP 1440
Db 1381 PAHSSAIGPVIGIILSLFVMGGVYFVQORVVQRYAGANGPPHYYVSGTTPHVPPLFIAP 1440
Qy 1441 GGSQHGPTGTIACGSKMSSVSLMGGRGVPLYDRNHVHTGASSSSSSSTKATLYPPIINP 1500
Db 1441 GGSQHGPTGTIACGSKMSSVSLMGGRGVPLYDRNHVHTGASSSSSSSTKATLYPPIINP 1500
Qy 1501 PPSPATDPSLYNMDMYSSNIPATARPVPIIIRGMAPPPTTPCSTDVCDSDYSASRWKAS 1560
Db 1501 PPSPATDPSLYNMDMYSSNIPATARPVPIIIRGMAPPPTTPCSTDVCDSDYSASRWKAS 1560
Qy 1561 KYLIDLNSDSDPYPPPTPHSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615
Db 1561 KYLIDLNSDSDPYPPPTPHSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615

RESULT 7
US-10-182-936A-3
; Sequence 3, Application US/10182936A
; Publication No. US20040038860A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Kristina M.
; APPLICANT: Anisowicz, Anthony
; APPLICANT: Bhat, Bheem
; APPLICANT: Damagnez, Veronique
; APPLICANT: Robinson, John
; APPLICANT: Yaworsky, Paul
; TITLE OF INVENTION: Reagents and Method for Modulating DKK-Mediated Interactions
; FILE REFERENCE: 032796-143
; CURRENT APPLICATION NUMBER: US/10/182,936A
; CURRENT FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: PCT/US02/15982
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: US 60/291,311
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: US 60/353,058
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: US 60/361,293
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 216
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1615
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-182-936A-3

Query Match 99.7%; Score 8715; DB 15; Length 1615;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1610; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MEAAPPPPPPLLLLLLLLLLALCCGPAPAAASPLLLFANRRDVLVDAGGVKLESTIVVS 60
Db 1 MEAAPPPPPPLLLLLLLLLLALCCGPAPAAASPLLLFANRRDVLVDAGGVKLESTIVVS 60
Qy 61 GLEDAADVDFQSKGAVYWTDVSEAIKQTYLNTGAAVQNVVIGLVSPDGLACDWGK 120
Db 61 GLEDAADVDFQSKGAVYWTDVSEAIKQTYLNTGAAVQNVVIGLVSPDGLACDWGK 120
Qy 121 KLYWTDSETRNRIEVANLNGTSRKVLFWQDLDPQRAIALDPAHGYMYWTWGTETPRIERAG 180
```

Db 121 KLYWTDSETRNIEVANLNGTSRKVLFWQDLDPKALDPAHYMYWTDWGETPRIERAG 180
Qy 181 MDGSTRKLIIVDSIIYWPNGLTIDLBQKLYWADAKLSFIHRANLDGSPQKVVESGLTTP 240
Db 181 MDGSTRKLIIVDSIIYWPNGLTIDLBQKLYWADAKLSFIHRANLDGSPQKVVESGLTTP 240
Qy 241 FALTISGDTLYWTDQTRSIHACNKTGGRKEIISALYSNMDIOVLQOEROPPFHTTCE 300
Db 241 FALTISGDTLYWTDQTRSIHACNKTGGRKEIISALYSNMDIOVLQOEROPPFHTTCE 300
Qy 301 EDNGCCHLCLSPSEPFYTCACPTGVQDNGRTCKAGAEVILLARRTLRISLDT 360
Db 301 EDNGCCHLCLSPSEPFYTCACPTGVQDNGRTCKAGAEVILLARRTLRISLDT 360
Qy 361 DFTDVLQVDDIRHAIADYDPLEGYVWTDDEVRAIRRAYIDGSGAOTLVNTEINDPDG 420
Db 361 DFTDVLQVDDIRHAIADYDPLEGYVWTDDEVRAIRRAYIDGSGAOTLVNTEINDPDG 420
Qy 421 IAVDMVARNLYWTDGTDRIEVTRNGTSRKILVSEDLDEPRAIALHPVMGLMYWTDWGE 480
Db 421 IAVDMVARNLYWTDGTDRIEVTRNGTSRKILVSEDLDEPRAIALHPVMGLMYWTDWGE 480
Qy 481 NPKIECANLDGQERRVLVNASLGNWGLALDQEGKLYWDAKTDKIEVINVDGFKRTL 540
Db 481 NPKIECANLDGQERRVLVNASLGNWGLALDQEGKLYWDAKTDKIEVINVDGFKRTL 540
Qy 541 LEDKLPHIFGFTLLGDFIYWTDQWRSIERVHKVAKASRDVIIDQLPDLMLKAVNAVW 600
Db 541 LEDKLPHIFGFTLLGDFIYWTDQWRSIERVHKVAKASRDVIIDQLPDLMLKAVNAVW 600
Qy 601 GTNPCADRNGGCHLCFFTPHATRCGPIGLELLSDMTKCIPEAFVFTSRAAHTRL 660
Db 601 GTNPCADRNGGCHLCFFTPHATRCGPIGLELLSDMTKCIPEAFVFTSRAAHTRL 660
Qy 661 ETNNNDVAIPITGVKEASALDFVSNHHIYWDVSLKTIISRAFMNGSSVEHVFGLDYP 720
Db 661 ETNNNDVAIPITGVKEASALDFVSNHHIYWDVSLKTIISRAFMNGSSVEHVFGLDYP 720
Qy 721 EGMADVMMKNIYADTCTNRIEVARLDGQROVLVWDLNPRSLALDPTKGYIYTW 780
Db 721 EGMADVMMKNIYADTCTNRIEVARLDGQROVLVWDLNPRSLALDPTKGYIYTW 780
Qy 781 GGPRIVAFMDGTCNMTLVKVGANDLTIDYADQRLYTDLDTNMISSNMLGQERV 840
Db 781 GGPRIVAFMDGTCNMTLVKVGANDLTIDYADQRLYTDLDTNMISSNMLGQERV 840
Qy 841 IADDLPHFPGLTQYSDIYTDWNLHSIERADKTSGRNRTLIQGLDFVMDILVFHSSRQ 900
Db 841 IADDLPHFPGLTQYSDIYTDWNLHSIERADKTSGRNRTLIQGLDFVMDILVFHSSRQ 900
Qy 901 DGLNDCHNNGCQGLCLAIPEGHRCGCASHYTLDPSSRNCSPPTTFLPSOKSAISRM 960
Db 901 DGLNDCHNNGCQGLCLAIPEGHRCGCASHYTLDPSSRNCSPPTTFLPSOKSAISRM 960
Qy 961 PDDQHSPLILPLHGLRNVAIDYDPLDKFIYWDGQRIKRAKDDGTQPFVLTSLSQG 1020
Db 961 PDDQHSPLILPLHGLRNVAIDYDPLDKFIYWDGQRIKRAKDDGTQPFVLTSLSQG 1020
Qy 1021 NPDRQPHDLSIDYISRTLFWTCEATNTINVRHLSGEAMGVLRGDRKPRAIIVNAERG 1080
Db 1021 NPDRQPHDLSIDYISRTLFWTCEATNTINVRHLSGEAMGVLRGDRKPRAIIVNAERG 1080
Qy 1081 LYFTNMQDRAAKIERAALDGTGEREVLFTTGLIRPVVALVVDNTLGLFWADADLKRISCD 1140
Db 1081 LYFTNMQDRAAKIERAALDGTGEREVLFTTGLIRPVVALVVDNTLGLFWADADLKRISCD 1140
Qy 1141 LSGANRLTLEDANIVQPLGLTILGKLYWIDRQOQMIERVEKTTGDKRTRIQRVAHLTG 1200
Db 1141 LSGANRLTLEDANIVQPLGLTILGKLYWIDRQOQMIERVEKTTGDKRTRIQRVAHLTG 1200
Qy 1201 IHAVEVSLEEFSAHPCARDNGGCSHCIAKDGCTPRCSCPVLHVLQNLITCCEPPTCS 1260
Db 1201 IHAVEVSLEEFSAHPCARDNGGCSHCIAKDGCTPRCSCPVLHVLQNLITCCEPPTCS 1260

RESULT 8

US-10-731-739-3
; Sequence 3, Application US/10731739
; Publication No. US20040176582A1
; GENERAL INFORMATION:
; APPLICANT: Carulli, John P.
; APPLICANT: Little, Randall D.
; APPLICANT: Recker, Robert R.
; APPLICANT: Johnson, Mark L.
; TITLE OF INVENTION: High bone mass gene of 11q13.3
; FILE OF INVENTION: 032796-013
; CURRENT APPLICATION NUMBER: US/10/731,739
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: US/09/544,398B
; PRIOR FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: US 09/229,319
; PRIOR FILING DATE: 1999-01-13
; PRIOR APPLICATION NUMBER: US 60/071,449
; PRIOR FILING DATE: 1998-01-13
; PRIOR APPLICATION NUMBER: US 60/105,511
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 641
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1615
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-731-739-3

Query Match 99.7%; Score 8715; DB 16; Length 1615;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1610; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MEAAPGPPWPLLLLLLLLLLALCCPAPAAASPLLLFANRRDRLVDAGGVKLESTIWS 60
Db 1 MEAAPGPPWPLLLLLLLLLLALCCPAPAAASPLLLFANRRDRLVDAGGVKLESTIWS 60
Qy 61 GLEDAADVDFQSKGAVYTDVSEEAIKOTYINOTGAAVQNVISGLVSPGLACDWVGK 120
Db 61 GLEDAADVDFQSKGAVYTDVSEEAIKOTYINOTGAAVQNVISGLVSPGLACDWVGK 120
Qy 121 KLYWTDSETRNIEVANLNGTSRKVLFWQDLDPKALDPAHYMYWTDWGETPRIERAG 180
Db 121 KLYWTDSETRNIEVANLNGTSRKVLFWQDLDPKALDPAHYMYWTDWGETPRIERAG 180
Qy 181 MDGSTRKLIIVDSIIYWPNGLTIDLBQKLYWADAKLSFIHRANLDGSPQKVVESGLTTP 240
Db 181 MDGSTRKLIIVDSIIYWPNGLTIDLBQKLYWADAKLSFIHRANLDGSPQKVVESGLTTP 240

Db 181 MDGSTRKIIIVDSDIYWPNGLTIDLEEQKLYWADAKLSFIHRANLDGSPRQKVVEGSLTHP 240
Qy 241 FALTLSGDTLTYWTDWQTRSIIHACNKRKTGGKREILSALYSMPDIQVLSQERQPPFHTRCE 300
Db 241 FALTLSGDTLTYWTDWQTRSIIHACNKRKTGGKREILSALYSMPDIQVLSQERQPPFHTRCE 300
Qy 301 EDNGGSHLCLLSSEPPYTCACPTGQVQDNGRTCKAGAEVLLARRTDLRRISLDT 360
Db 301 EDNGGSHLCLLSSEPPYTCACPTGQVQDNGRTCKAGAEVLLARRTDLRRISLDT 360
Qy 361 DFTDVLQVDDIRAIADIDYDPLSGYVYVTDDEVRATRRAYLDGSGAQTLVNTINPDG 420
Db 361 DFTDVLQVDDIRAIADIDYDPLSGYVYVTDDEVRATRRAYLDGSGAQTLVNTINPDG 420
Qy 421 IAVDWARNLYWTDGTDRIEVTRNGTSRKILVSEDLDEPRALALHPVGLMTWTDGE 480
Db 421 IAVDWARNLYWTDGTDRIEVTRNGTSRKILVSEDLDEPRALALHPVGLMTWTDGE 480
Qy 481 NPKIECANLDQOERRVLVNASLGNPGLALDLQEGKLYWGDAKTKIEVINVDGTRKRTL 540
Db 481 NPKIECANLDQOERRVLVNASLGNPGLALDLQEGKLYWGDAKTKIEVINVDGTRKRTL 540
Qy 541 LEDKLPHFIFGTLLGDFIYVTDWQRRSIEVHVKKASRDVITDQPLMGLKAVNAKV 600
Db 541 LEDKLPHFIFGTLLGDFIYVTDWQRRSIEVHVKKASRDVITDQPLMGLKAVNAKV 600
Qy 601 GTNPACDRNGGCSHLCPFTPHATRCGPIGLELLSDMKTCIVPEAFVFTSRAAHRISL 660
Db 601 GTNPACDRNGGCSHLCPFTPHATRCGPIGLELLSDMKTCIVPEAFVFTSRAAHRISL 660
Qy 661 ETNNNDVAIPLTGVKEASALDFOVSNHHIYWTDVSLKTIISAFVNGSSVEHVVEGLDYP 720
Db 661 ETNNNDVAIPLTGVKEASALDFOVSNHHIYWTDVSLKTIISAFVNGSSVEHVVEGLDYP 720
Qy 721 EGMADVAMGNLYWADGTNRIEVARLDGQROVLVWRDLNPRSLALDPTKGIIYVTEW 780
Db 721 EGMADVAMGNLYWADGTNRIEVARLDGQROVLVWRDLNPRSLALDPTKGIIYVTEW 780
Qy 781 GKKPRIVRAFMDGNTCMVLVDKVRANDLTIDYADQRLYWTDLTNMIESNMGLQSERVV 840
Db 781 GKKPRIVRAFMDGNTCMVLVDKVRANDLTIDYADQRLYWTDLTNMIESNMGLQSERVV 840
Qy 841 IADLPHPPGLTQYSDIYIYVTDWNLHSEIRADKTSGRNRTLIOGHLDVMDILVFHSSRQ 900
Db 841 IADLPHPPGLTQYSDIYIYVTDWNLHSEIRADKTSGRNRTLIOGHLDVMDILVFHSSRQ 900
Qy 901 DGLNDCMHNNGCQQLCIAIPGHRCCGASHYTLDPSSRNCSPPTTFLFSQKSAISRM 960
Db 901 DGLNDCMHNNGCQQLCIAIPGHRCCGASHYTLDPSSRNCSPPTTFLFSQKSAISRM 960
Qy 961 PDDQHSPLILPLHGLRNVAIDYDPLDKFYWYDGRQNIKRAKDDGTQPPVLTLSQGG 1020
Db 961 PDDQHSPLILPLHGLRNVAIDYDPLDKFYWYDGRQNIKRAKDDGTQPPVLTLSQGG 1020
Qy 1021 NPDRQPHDLSDIYSRTLFWTCEATNTINVHRLSGEAMGVVLRGDRDKPRAIVVNAERG 1080
Db 1021 NPDRQPHDLSDIYSRTLFWTCEATNTINVHRLSGEAMGVVLRGDRDKPRAIVVNAERG 1080
Qy 1081 LYFTNMQDRAKIERAALDGTREVLFTTGLIRPVALVWNTIQLFWADLKRIESCD 1140
Db 1081 LYFTNMQDRAKIERAALDGTREVLFTTGLIRPVALVWNTIQLFWADLKRIESCD 1140
Qy 1141 LSGANRLTLTDANIVQPLGLTILGKHYWIDRQOQMTERVEKTTGDKRTRIQGRVAHJTG 1200
Db 1141 LSGANRLTLTDANIVQPLGLTILGKHYWIDRQOQMTERVEKTTGDKRTRIQGRVAHJTG 1200
Qy 1201 IHAVEVSLSEFSAHPCARDNGGCSHTCIAKGDGTPRCSCPVLHVLQNLTLTCEPTCS 1260
Db 1201 IHAVEVSLSEFSAHPCARDNGGCSHTCIAKGDGTPRCSCPVLHVLQNLTLTCEPTCS 1260
Qy 1261 PDQACATGIDICIPGAWRCDGPFECDDQSDDEGCPVCSAAQPPCARGQCVDLRLRCDGE 1320
Db 1261 PDQACATGIDICIPGAWRCDGPFECDDQSDDEGCPVCSAAQPPCARGQCVDLRLRCDGE 1320

Qy 1321 ADCODRDEADCDALCLPNQFRCASGQCVLIKQCCDSDPDICIDGSDLMCBITTKPPSDS 1380
Db 1321 ADCODRDEADCDALCLPNQFRCASGQCVLIKQCCDSDPDICIDGSDLMCBITTKPPSDS 1380
Qy 1381 PAHSSAIGPVGIIILSLFVMGVYFVQVVCQRYAGANGPPPHYVSGTDPHVLNFIAP 1440
Db 1381 PAHSSAIGPVGIIILSLFVMGVYFVQVVCQRYAGANGPPPHYVSGTDPHVLNFIAP 1440
Qy 1441 GGSQHGPTGTACGKSMSSVSLMGGRGVPLYDRNHVGTGASSSSSSSTKATLYPPIINP 1500
Db 1441 GGSQHGPTGTACGKSMSSVSLMGGRGVPLYDRNHVGTGASSSSSSSTKATLYPPIINP 1500
Qy 1501 PPSPATDPSLYNMDPFYSSNIPATAPYRPIYIIRGMAPTTTSCSTDVCDSDYSASRWKAS 1560
Db 1501 PPSPATDPSLYNMDPFYSSNIPATAPYRPIYIIRGMAPTTTSCSTDVCDSDYSASRWKAS 1560
Qy 1561 KYLDLNSDSDPYPPPTPHSQYLSAEDSCPPSPTATERSYFHLFPPPPSPCTDSS 1615
Db 1561 KYLDLNSDSDPYPPPTPHSQYLSAEDSCPPSPTATERSYFHLFPPPPSPCTDSS 1615

RESULT 9

US-10-374-979-4
; Sequence 4, Application US/10374979
; Publication No. US2003021979A1
; GENERAL INFORMATION:
; APPLICANT: John P. Carulli et al.
; TITLE OF INVENTION: THE HIGH BONE MASS GENE OF 11q13.3
; FILE REFERENCE: 032796-021
; CURRENT APPLICATION NUMBER: US/10/374,979
; CURRENT FILING DATE: 2003-03-04
; PRIOR APPLICATION NUMBER: US 09/544,398
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 09/543,771
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 09/229,319
; PRIOR FILING DATE: 1999-01-13
; PRIOR APPLICATION NUMBER: US 60/071,449
; PRIOR FILING DATE: 1998-01-13
; PRIOR APPLICATION NUMBER: US 60/105,511
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 109
; SEQ ID NO 4
; LENGTH: 1615
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-374-979-4

Query Match 99.6%; Score 8706; DB 15; Length 1615;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1609; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MEAAPGPPWPLLLLLLLLLLALCGCPAPAAAPLFFANRRDVRLLVADAGGVKLESTIVVS 60
Db 1 MEAAPGPPWPLLLLLLLLLLALCGCPAPAAAPLFFANRRDVRLLVADAGGVKLESTIVVS 60
Qy 61 GLEDAAVDQFSKGVYVTDVSEAIKOTVLTGTAQVNVISGLVSPGLACDWWGK 120
Db 61 GLEDAAVDQFSKGVYVTDVSEAIKOTVLTGTAQVNVISGLVSPGLACDWWGK 120
Qy 121 KLYWTDSETRNRIEVANLNGTSRKVLFWQDLDQPRALDPAHGYMYWTDGTPRIERAG 180
Db 121 KLYWTDSETRNRIEVANLNGTSRKVLFWQDLDQPRALDPAHGYMYWTDGTPRIERAG 180
Qy 181 MDGSTRKIIIVDSDIYWPNGLTIDLEEQKLYWADAKLSFIHRANLDGSPRQKVVEGSLTHP 240
Db 181 MDGSTRKIIIVDSDIYWPNGLTIDLEEQKLYWADAKLSFIHRANLDGSPRQKVVEGSLTHP 240
Qy 241 FALTLSGDTLTYWTDWQTRSIIHACNKRKTGGKREILSALYSMPDIQVLSQERQPPFHTRCE 300
Db 241 FALTLSGDTLTYWTDWQTRSIIHACNKRKTGGKREILSALYSMPDIQVLSQERQPPFHTRCE 300

```
QY 301 EDNGGCSHLCLLSPSEPFYTCACPTGVOLQDNRTCKAGABEVLLARRTDLRLSLDTP 360
Db 301 EDNGGWSHLCLLSPSEPFYTCACPTGVQMDNGRTCKAGABEVLLARRTDLRLSLDTP 360
QY 361 DFTDVLQVDDIRHAIADIDPLEGIVYVWTDDEVRARRAYLDGSGAQTLVNTEINDPDG 420
Db 361 DFTDVLQVDDIRHAIADIDPLEGIVYVWTDDEVRARRAYLDGSGAQTLVNTEINDPDG 420
QY 421 IAVDVARNLYWTDGTDRIEVTRLANGSRKLLVSEDLDEPRRAIALHPVMGLMYWTDWGE 480
Db 421 IAVDVARNLYWTDGTDRIEVTRLANGSRKLLVSEDLDEPRRAIALHPVMGLMYWTDWGE 480
QY 481 NPKIECANLQGERVNLVNASLWPNGLALDQEGKLYWGAKTDKIEVINVDGTKRRTL 540
Db 481 NPKIECANLQGERVNLVNASLWPNGLALDQEGKLYWGAKTDKIEVINVDGTKRRTL 540
QY 541 LEDKLPHIFGFTLLGDFIYVTDWQRRSIRVHKVKSARDVIIDQLPDLMLGLKANVAKV 600
Db 541 LEDKLPHIFGFTLLGDFIYVTDWQRRSIRVHKVKSARDVIIDQLPDLMLGLKANVAKV 600
QY 601 GTNPCADRNGCASHLCFTPHATRCGCPIGLELLSDMTKTCIVPEAFVFTSRAAIIHRLS 660
Db 601 GTNPCADRNGCASHLCFTPHATRCGCPIGLELLSDMTKTCIVPEAFVFTSRAAIIHRLS 660
QY 661 ETNNNDVAIPLTVKESALDQFVSNHHIYVTDVSLKTIISRAFNGSSVEHVVERGLDYP 720
Db 661 ETNNNDVAIPLTVKESALDQFVSNHHIYVTDVSLKTIISRAFNGSSVEHVVERGLDYP 720
QY 721 EGMVDMWGNLYWADTGTNRIEVARLDQFQVRLVWRDLNPRSLALDPTKGYIYTWEM 780
Db 721 EGMVDMWGNLYWADTGTNRIEVARLDQFQVRLVWRDLNPRSLALDPTKGYIYTWEM 780
QY 781 GSKPRIVAFWGTNCMTLVKVGANDLTIDYADORLYWTDLTNTMIESNMGLQERVV 840
Db 781 GSKPRIVAFWGTNCMTLVKVGANDLTIDYADORLYWTDLTNTMIESNMGLQERVV 840
QY 841 IADDLPHFPGFTQYSDIYVTDWNLHLSIERADKTSGRNRTLIQGHLDVWMDILVPHSRQ 900
Db 841 IADDLPHFPGFTQYSDIYVTDWNLHLSIERADKTSGRNRTLIQGHLDVWMDILVPHSRQ 900
QY 901 DGLNDCMHNQCQGLCLAIIPGHRCCGASHYTLDPSSRNCSPPTTFLFFSKSAISMI 960
Db 901 DGLNDCMHNQCQGLCLAIIPGHRCCGASHYTLDPSSRNCSPPTTFLFFSKSAISMI 960
QY 961 PDDQHSPLIILPLHGLRNKALIDVPLDKFYVWDGQNIKRAKDDGTQPFVLTSLSQG 1020
Db 961 PDDQHSPLIILPLHGLRNKALIDVPLDKFYVWDGQNIKRAKDDGTQPFVLTSLSQG 1020
QY 1021 NPDRQPHDLSIDIYSRTLFWTCEATNTINVHLSGEAMGVVLRGDRDKPRAIVVNAERG 1080
Db 1021 NPDRQPHDLSIDIYSRTLFWTCEATNTINVHLSGEAMGVVLRGDRDKPRAIVVNAERG 1080
QY 1081 LYFTNMQDRAAKIERAALDGTREVEVFTTGLIRPVALVVDNTLQKLFWVDADLKRIBSCD 1140
Db 1081 LYFTNMQDRAAKIERAALDGTREVEVFTTGLIRPVALVVDNTLQKLFWVDADLKRIBSCD 1140
QY 1141 LSGANRLTLEDANIVQPLGLTILGKHLWIDRQOQMIEERVEKTTGDKRTRIQGRVAHLTG 1200
Db 1141 LSGANRLTLEDANIVQPLGLTILGKHLWIDRQOQMIEERVEKTTGDKRTRIQGRVAHLTG 1200
QY 1201 IHAVEVSLEEFSAHPCARDNGGCSHICIAKGDGTPRCSPVHLVLLQNLLTCGEPTCS 1260
Db 1201 IHAVEVSLEEFSAHPCARDNGGCSHICIAKGDGTPRCSPVHLVLLQNLLTCGEPTCS 1260
QY 1261 PDQACATGEIDCIPGAWRCDGPFECDDQSDSECPVCSAAQFFCARGQCQVDLRLCDGE 1320
Db 1261 PDQACATGEIDCIPGAWRCDGPFECDDQSDSECPVCSAAQFFCARGQCQVDLRLCDGE 1320
QY 1321 ADCQDRSDEADCDAILCPNPRCAGQCVLIKQCDGFPDCCIDGSDLMCEITKPPSDDS 1380
Db 1321 ADCQDRSDEADCDAILCPNPRCAGQCVLIKQCDGFPDCCIDGSDLMCEITKPPSDDS 1380
QY 1381 PAHSSAIGPVIIGIILSLFVMGGVYFVCORVVCQRYAGANGFPFHEVSGTTPHPLNFIAP 1440
```

```
Db 1381 PAHSSAIGPVIIGIILSLFVMGGVYFVCORVVCQRYAGANGFPFHEVSGTTPHPLNFIAP 1440
QY 1441 GGSQHGPFTGIACGKSMSSVSLMGGRGVPLYDRNHVVTGASSSSSSSTKATLYPPIILNP 1500
Db 1441 GGSQHGPFTGIACGKSMSSVSLMGGRGVPLYDRNHVVTGASSSSSSSTKATLYPPIILNP 1500
QY 1501 PPSPATDPSLNMDFYSSNIPATARPVRPIYIRGMAPPPTTSCSTDVCDSDYSASRWKAS 1560
Db 1501 PPSPATDPSLNMDFYSSNIPATARPVRPIYIRGMAPPPTTSCSTDVCDSDYSASRWKAS 1560
QY 1561 KYIYDLNLSDSPPYPPPTPHSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615
Db 1561 KYIYDLNLSDSPPYPPPTPHSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615

RESULT 10
US-10-182-936A-4
; Sequence 4, Application US/10182936A
; Publication No. US20040038860A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Kristina M.
; APPLICANT: Anisowicz, Anthony
; APPLICANT: Bhat, Bheem
; APPLICANT: Damaghez, Veronique
; APPLICANT: Robinson, John
; APPLICANT: Yaworsky, Paul
; TITLE OF INVENTION: Reagents and Method for Modulating DKK-Mediated Interactions
; FILE REFERENCE: 032796-143
; CURRENT APPLICATION NUMBER: US/10/182,936A
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: PCT/US02/15982
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: US 60/291,311
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: US 60/353,058
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: US 60/361,293
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 216
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1615
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-182-936A-4

Query Match 99.6%; Score 8706; DB 15; Length 1615;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1609; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 MEAAPGPPPLLLLLLLLLLALCGCPAPAAASPLLLFANRRDRLVLDAGGVKLESTIVS 60
Db 1 MEAAPGPPPLLLLLLLLLLALCGCPAPAAASPLLLFANRRDRLVLDAGGVKLESTIVS 60
QY 61 GLEDAADVDFQSKGAVYVTDVSEEAIKQYLNQTAGVONVVTSGLVSPDGLACDWGK 120
Db 61 GLEDAADVDFQSKGAVYVTDVSEEAIKQYLNQTAGVONVVTSGLVSPDGLACDWGK 120
QY 121 KLYWTDSETNRIEVANLNGTSRKVLFWQDLDQPRALDPAHGYWYVTDWGETPRIERAG 180
Db 121 KLYWTDSETNRIEVANLNGTSRKVLFWQDLDQPRALDPAHGYWYVTDWGETPRIERAG 180
QY 181 MDGSTRKIIYVDSIYWPNGLTIDLEQKLYWADAKLSFIHRANLDGSRQKRVGSLTTP 240
Db 181 MDGSTRKIIYVDSIYWPNGLTIDLEQKLYWADAKLSFIHRANLDGSRQKRVGSLTTP 240
QY 241 FALTLSGDTLYWTDWQTRSIIHACNKRGTGKRKEILSALYSPMDIQVLSQROPPFHTRCE 300
Db 241 FALTLSGDTLYWTDWQTRSIIHACNKRGTGKRKEILSALYSPMDIQVLSQROPPFHTRCE 300
QY 301 EDNGGCSHLCLLSPSEPFYTCACPTGVOLQDNRTCKAGABEVLLARRTDLRLSLDTP 360
```

Db 301 EDNGGWSHLCILLSPEPFTYACPTGVQMDNGRTCKAGAEVLLARRTDLRLSLDTP 360
Qy 361 DFTDVLQVDDIRAHAIADYDPLEGYVYWTDEVRAIRRAYLDGSGAOTLVNTEINDPDG 420
Db 361 DFTDVLQVDDIRAHAIADYDPLEGYVYWTDEVRAIRRAYLDGSGAOTLVNTEINDPDG 420
Qy 421 IAVDWARNLYWTGTDRIEVTRNGTSRKILVSEDLDEPRAIALHPVGMGLMYWTWGE 480
Db 421 IAVDWARNLYWTGTDRIEVTRNGTSRKILVSEDLDEPRAIALHPVGMGLMYWTWGE 480
Qy 481 NPKTECANLDCQERRVYNASLGNPGLALDQEGKLYWDAKTDKIEVINVDGTKRTL 540
Db 481 NPKTECANLDCQERRVYNASLGNPGLALDQEGKLYWDAKTDKIEVINVDGTKRTL 540
Qy 541 LEDKLPHIFGFTLLGDFIYWTDMQRRSIEVHVKKASRDVLIIDQLPDLMLKAVNAVAV 600
Db 541 LEDKLPHIFGFTLLGDFIYWTDMQRRSIEVHVKKASRDVLIIDQLPDLMLKAVNAVAV 600
Qy 601 GTNFCADRNGGCSHLCFFTPHATRCGCPIGLELLSDMKTCIVPEAFVFTSRAAHRISL 660
Db 601 GTNFCADRNGGCSHLCFFTPHATRCGCPIGLELLSDMKTCIVPEAFVFTSRAAHRISL 660
Qy 661 ETNNNDVAIPLTGVKEASALDFVSNHHIYWTDSVLSKTIISAFMNGSSVEHVFEGLDYP 720
Db 661 ETNNNDVAIPLTGVKEASALDFVSNHHIYWTDSVLSKTIISAFMNGSSVEHVFEGLDYP 720
Qy 721 EGMAVDMGKNLYWADTGTNRIEVARLDGQFQVLRWDLNPRSLALDPTKGYIYTW 780
Db 721 EGMAVDMGKNLYWADTGTNRIEVARLDGQFQVLRWDLNPRSLALDPTKGYIYTW 780
Qy 781 GKGPRIVRAFMGDMGNTLVKVGANDLTIDYADQRLYWTDLNMISSNMLGQERVV 840
Db 781 GKGPRIVRAFMGDMGNTLVKVGANDLTIDYADQRLYWTDLNMISSNMLGQERVV 840
Qy 841 IADLPHFPGLTQYSDIYIYTDNWLHSIERADKTSGRNRTLIQGLHDFVMDILVFHSRQ 900
Db 841 IADLPHFPGLTQYSDIYIYTDNWLHSIERADKTSGRNRTLIQGLHDFVMDILVFHSRQ 900
Qy 901 DGLNDCHNNGCQGLCLAIPEGHRCGCASHYTLDPSSRNCSPPTTFLFSQKSAISRM 960
Db 901 DGLNDCHNNGCQGLCLAIPEGHRCGCASHYTLDPSSRNCSPPTTFLFSQKSAISRM 960
Qy 961 PDDQHSPLILPLHGLRNVAIDYDPLDKFYIYWDGQNIKRAKDXDGTQPPVLTSLSQG 1020
Db 961 PDDQHSPLILPLHGLRNVAIDYDPLDKFYIYWDGQNIKRAKDXDGTQPPVLTSLSQG 1020
Qy 1021 NPDQPHDLSIDIYSRTLFWTCEATNTINVRHLSGEAMGVVLRGDRKPRAIIVNAERG 1080
Db 1021 NPDQPHDLSIDIYSRTLFWTCEATNTINVRHLSGEAMGVVLRGDRKPRAIIVNAERG 1080
Qy 1081 LYFTNMDRAAKIERAALDGTREVLFTTGLIRPVVALVVDNTLGLFWVDADLKRIBSCD 1140
Db 1081 LYFTNMDRAAKIERAALDGTREVLFTTGLIRPVVALVVDNTLGLFWVDADLKRIBSCD 1140
Qy 1141 LSGANRLTLEDANIVQPLGLTILGKHLWIYDROQQMIERVEKTTGDKXRTIQRVAHLTG 1200
Db 1141 LSGANRLTLEDANIVQPLGLTILGKHLWIYDROQQMIERVEKTTGDKXRTIQRVAHLTG 1200
Qy 1201 IHAVEESLEEFSAHPCARDNGGCSHICIAKGDGTPRCSPVHLVLQNLLTCGEPPTCS 1260
Db 1201 IHAVEESLEEFSAHPCARDNGGCSHICIAKGDGTPRCSPVHLVLQNLLTCGEPPTCS 1260
Qy 1261 PDQFACATGEIDCIPGAWRCGDPPECDDQSDDEGCPVCSAQQPCARGCQVDLRLCDGE 1320
Db 1261 PDQFACATGEIDCIPGAWRCGDPPECDDQSDDEGCPVCSAQQPCARGCQVDLRLCDGE 1320
Qy 1321 ADCQDRSDEADCAICLPNQPRCASGQCVLIQKQCDSPDCIDGSDDELMECEITPPSDDS 1380
Db 1321 ADCQDRSDEADCAICLPNQPRCASGQCVLIQKQCDSPDCIDGSDDELMECEITPPSDDS 1380
Qy 1381 PAHSSAIGPVTGIILSLFVPMGGVYFVQVQVQRYAGANGFPPEHYVSGTTPHPLNFIAP 1440
Db 1381 PAHSSAIGPVTGIILSLFVPMGGVYFVQVQVQRYAGANGFPPEHYVSGTTPHPLNFIAP 1440

Qy 1441 GGSQHGPFTGIACGKSMSSVSLMGGRGVPLYDRNHVHTGASSSSSTKATLYPPIILNP 1500
Db 1441 GGSQHGPFTGIACGKSMSSVSLMGGRGVPLYDRNHVHTGASSSSSTKATLYPPIILNP 1500
Qy 1501 PPSPATDPSLYNMDMFYSSNIPATAPYRPIYIIRGMAPPPTTTCSTDVCDSDYSASRWKAS 1560
Db 1501 PPSPATDPSLYNMDMFYSSNIPATAPYRPIYIIRGMAPPPTTTCSTDVCDSDYSASRWKAS 1560
Qy 1561 KYLDLANSDSDPYPPPTPHSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615
Db 1561 KYLDLANSDSDPYPPPTPHSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615

RESULT 11
US-10-731-739-4
; Sequence 4, Application US/10731739
; Publication No. US20040176582A1
; GENERAL INFORMATION:
; APPLICANT: Carulli, John P.
; APPLICANT: Little, Randall D.
; APPLICANT: Recker, Robert R.
; APPLICANT: Johnson, Mark L.
; TITLE OF INVENTION: High bone mass gene of 11q13.3
; FILE REFERENCE: 032796-013
; CURRENT APPLICATION NUMBER: US/10/731,739
; PRIOR FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: US/09/544,398B
; PRIOR FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: US 09/229,319
; PRIOR FILING DATE: 1999-01-13
; PRIOR APPLICATION NUMBER: US 60/071,449
; PRIOR FILING DATE: 1998-01-13
; PRIOR APPLICATION NUMBER: US 60/105,511
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 641
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1615
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-731-739-4

Query Match 99.6%; Score 8706; DB 16; Length 1615;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1609; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MEAAPGPPWPLLLLLLLALCCGPAPAAASPLLLFANRRDRLVDAGGVKLESTIVVS 60
Db 1 MEAAPGPPWPLLLLLLLALCCGPAPAAASPLLLFANRRDRLVDAGGVKLESTIVVS 60
Qy 61 GLEDAAAVDFQPSKGVYWTDVSEAIKQTYLNOTGAQVNVISGLVSPGLACDWGK 120
Db 61 GLEDAAAVDFQPSKGVYWTDVSEAIKQTYLNOTGAQVNVISGLVSPGLACDWGK 120
Qy 121 KLYWTDSETRNRIEVANLNGTSRKVLFWQDLDDQPAIALDPAHYMYWTDWGETPRIERAG 180
Db 121 KLYWTDSETRNRIEVANLNGTSRKVLFWQDLDDQPAIALDPAHYMYWTDWGETPRIERAG 180
Qy 181 MDGSTRKIIIVDSDIYWPNGLTIDLEEQLYWADAKLSFIHRANLDGSRQKVGESLTHP 240
Db 181 MDGSTRKIIIVDSDIYWPNGLTIDLEEQLYWADAKLSFIHRANLDGSRQKVGESLTHP 240
Qy 241 FALTLSGDTLYWTDWQTRSIIHACNRTGGKKEKILSALYSPMDIQLVLSQRPFFHTRCE 300
Db 241 FALTLSGDTLYWTDWQTRSIIHACNRTGGKKEKILSALYSPMDIQLVLSQRPFFHTRCE 300
Qy 301 EDNGGSHLCILLSPEPFTYACPTGVQMDNGRTCKAGAEVLLARRTDLRLSLDTP 360
Db 301 EDNGGSHLCILLSPEPFTYACPTGVQMDNGRTCKAGAEVLLARRTDLRLSLDTP 360
Qy 361 DFTDVLQVDDIRAHAIADYDPLEGYVYWTDEVRAIRRAYLDGSGAOTLVNTEINDPDG 420
Db 361 DFTDVLQVDDIRAHAIADYDPLEGYVYWTDEVRAIRRAYLDGSGAOTLVNTEINDPDG 420

Db 361 DFTDIVLQVDDIRHAIADYDPLEGVVYVWTDDEVRAIRRAYLDGSGAQTLYNTEINDPDG 420
Qy 421 IAVDWARNLYWTDGTDRIEVTRNGTSRKILVSEDLDEPRALIAHPVGLMYWTDWGE 480
Db 421 IAVDWARNLYWTDGTDRIEVTRNGTSRKILVSEDLDEPRALIAHPVGLMYWTDWGE 480
Qy 481 NPKIECANLQGERVLVNASIGWPNGLALDLQEGKLYWGDAKTDKIEVINVDGFKRTL 540
Db 481 NPKIECANLQGERVLVNASIGWPNGLALDLQEGKLYWGDAKTDKIEVINVDGFKRTL 540
Qy 541 LEDKLPHIFGFTLLGDFIYWTDMQRRSIERVHKVKSARDVIIDQLPDLMLGLKANVAVKV 600
Db 541 LEDKLPHIFGFTLLGDFIYWTDMQRRSIERVHKVKSARDVIIDQLPDLMLGLKANVAVKV 600
Qy 601 GTNPCADRNGGSHLCFTPHATRCGPIGIELLSDMKTCIVPEAFVFTSRAAIHRISL 660
Db 601 GTNPCADRNGGSHLCFTPHATRCGPIGIELLSDMKTCIVPEAFVFTSRAAIHRISL 660
Qy 661 ETNNNDVAIPLTGVKEASALDPDVSNHHIYWTDSLKTSIRAFMNGSSVEHVVEGLDYP 720
Db 661 ETNNNDVAIPLTGVKEASALDPDVSNHHIYWTDSLKTSIRAFMNGSSVEHVVEGLDYP 720
Qy 721 EGMVDMWGNKLYWADTGTNRIEVARLDGQFQVLRVWRDLNPRSLALDPTKGYIYWTEW 780
Db 721 EGMVDMWGNKLYWADTGTNRIEVARLDGQFQVLRVWRDLNPRSLALDPTKGYIYWTEW 780
Qy 781 GKGPRIVRAFDGTCNMTLVKVGANDLTIDYADQRLYWTDLDTNMISSNMLGOERVV 840
Db 781 GKGPRIVRAFDGTCNMTLVKVGANDLTIDYADQRLYWTDLDTNMISSNMLGOERVV 840
Qy 841 IADDLPHFPFGLTQYSDIYIYWTDMNLHSIERADKTSGRNRTLIQGHLPFVMDILVPHSSRQ 900
Db 841 IADDLPHFPFGLTQYSDIYIYWTDMNLHSIERADKTSGRNRTLIQGHLPFVMDILVPHSSRQ 900
Qy 901 DGLNDGMHNGCGQQLCIAIPGHRGCGCASHYTLDPSSRNCSPPTTFLFQSKAISRM 960
Db 901 DGLNDGMHNGCGQQLCIAIPGHRGCGCASHYTLDPSSRNCSPPTTFLFQSKAISRM 960
Qy 961 PDDQHSPLIILPLHGLRNVAIDYDPLDKFIYWDGRONIKRAKDDGTQPPVLTSLSGQ 1020
Db 961 PDDQHSPLIILPLHGLRNVAIDYDPLDKFIYWDGRONIKRAKDDGTQPPVLTSLSGQ 1020
Qy 1021 NPDRQPHDLSIDIYSRTLFWTCEATNTINVRHLSGEAMGVVLRGDRDKPRAIVNAERG 1080
Db 1021 NPDRQPHDLSIDIYSRTLFWTCEATNTINVRHLSGEAMGVVLRGDRDKPRAIVNAERG 1080
Qy 1081 LYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVVAVVDNTLGLFWVDADLKRIBSCD 1140
Db 1081 LYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVVAVVDNTLGLFWVDADLKRIBSCD 1140
Qy 1141 LSGANRLTLEDANIVQPLGLTLGKHLWIDRQOQMIERVEKTTGDKETRIQGRVAHLTG 1200
Db 1141 LSGANRLTLEDANIVQPLGLTLGKHLWIDRQOQMIERVEKTTGDKETRIQGRVAHLTG 1200
Qy 1201 IHAVEEVSLEEFSAHPCARDNGGSHICIAKAGDGTGPRCSFVHLVLLQNLITCGEPPTCS 1260
Db 1201 IHAVEEVSLEEFSAHPCARDNGGSHICIAKAGDGTGPRCSFVHLVLLQNLITCGEPPTCS 1260
Qy 1261 PDQFACATGEIDCIPGAWRCDGPFECDDQSBEGCPVCSAAQFFCARGQCVDLRLRCDGE 1320
Db 1261 PDQFACATGEIDCIPGAWRCDGPFECDDQSBEGCPVCSAAQFFCARGQCVDLRLRCDGE 1320
Qy 1321 ADCQDRSEADCDACILPNQFRCSAGQCVLLKQCDSPPDCIDGSDLMCEITKPPSDDS 1380
Db 1321 ADCQDRSEADCDACILPNQFRCSAGQCVLLKQCDSPPDCIDGSDLMCEITKPPSDDS 1380
Qy 1381 PAHSSAIGPVGIIILSLFVWGGVYFVQCVVQRYAGANGFPFHEYSGTGHVPLNFIAP 1440
Db 1381 PAHSSAIGPVGIIILSLFVWGGVYFVQCVVQRYAGANGFPFHEYSGTGHVPLNFIAP 1440
Qy 1441 GGSQHGFTGTACGKSMSSVSLMGGRGVPLYDRNHVTGASSSSSTKATLYPPILNP 1500
Db 1441 GGSQHGFTGTACGKSMSSVSLMGGRGVPLYDRNHVTGASSSSSTKATLYPPILNP 1500

Qy 1501 PPSPATDPSLNMDFYSSNIPATAPRYRPIYIRGMAPPPTPCSTDVCDSDYSASRWKAS 1560
Db 1501 PPSPATDPSLNMDFYSSNIPATAPRYRPIYIRGMAPPPTPCSTDVCDSDYSASRWKAS 1560
Qy 1561 KYIYDLNLSDSPPYPPPTPHSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615
Db 1561 KYIYDLNLSDSPPYPPPTPHSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615

RESULT 12

US-10-464-368-81
; Sequence 81, Application US/10464368
; Publication No. US20040023356A1
; GENERAL INFORMATION:
; APPLICANT: Krumlauf, Robb
; APPLICANT: Ellies, Debra
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR REGULATING BONE DEPOSITION
; FILE REFERENCE: 40716-IP-017
; CURRENT APPLICATION NUMBER: US/10/464,368
; PRIOR FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: 60/388,970
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 81
; LENGTH: 1611
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-10-464-368-81

Query Match 99.6%; Score 8705; DB 15; Length 1611;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 1611; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

Qy 1 MEAAPPPPPPLLLLLLLLLLALCGCPAPAAASPLLLFANRRDRVLVDAAGVKLESTIVVS 60
Db 1 MEAAPPPPPPLLLLLLLLLLALCGCPAPAAASPLLLFANRRDRVLVDAAGVKLESTIVVS 60
Qy 61 GLEDAADAVQFSGKGVYVTDVSEBIAKQTYLNTGAAVQNVVIGSLVSPGLACDWVGK 120
Db 61 GLEDAADAVQFSGKGVYVTDVSEBIAKQTYLNTGAAVQNVVIGSLVSPGLACDWVGK 120
Qy 121 KLYWTDSETRNRIEVANLNGTSRKVLFWQDLDDOPRAIADPAHGYWYTDWGETRIERAG 180
Db 121 KLYWTDSETRNRIEVANLNGTSRKVLFWQDLDDOPRAIADPAHGYWYTDWGETRIERAG 180
Qy 181 MDGSTRKIIIVDSDIYWPNGLTIDLEEQKLYWADAKLSFIHRANLDGSPQKVVESGLTTP 240
Db 181 MDGSTRKIIIVDSDIYWPNGLTIDLEEQKLYWADAKLSFIHRANLDGSPQKVVESGLTTP 240
Qy 241 FALTLSGDTLYWTDWQTRSIAHCKRTGGKKEILSALYSPMDIQLVSOERQPFHTRCE 300
Db 241 FALTLSGDTLYWTDWQTRSIAHCKRTGGKKEILSALYSPMDIQLVSOERQPFHTRCE 300
Qy 301 EDNGGCHLCLLSPSEFPYTCACPTGVLQDNGRTCKAGAEVLLARRTDLRISLDTP 360
Db 301 EDNGGCHLCLLSPSEFPYTCACPTGVLQDNGRTCKAGAEVLLARRTDLRISLDTP 360
Qy 361 DFTDIVLQVDDIRHAIADYDPLEGVVYVWTDDEVRAIRRAYLDGSGAQTLYNTEINDPDG 420
Db 361 DFTDIVLQVDDIRHAIADYDPLEGVVYVWTDDEVRAIRRAYLDGSGAQTLYNTEINDPDG 420
Qy 421 IAVDWARNLYWTDGTDRIEVTRNGTSRKILVSEDLDEPRALIAHPVGLMYWTDWGE 480
Db 421 IAVDWARNLYWTDGTDRIEVTRNGTSRKILVSEDLDEPRALIAHPVGLMYWTDWGE 480
Qy 481 NPKIECANLQGERVLVNASIGWPNGLALDLQEGKLYWGDAKTDKIEVINVDGFKRTL 540
Db 481 NPKIECANLQGERVLVNASIGWPNGLALDLQEGKLYWGDAKTDKIEVINVDGFKRTL 540
Qy 541 LEDKLPHIFGFTLLGDFIYWTDMQRRSIERVHKVKSARDVIIDQLPDLMLGLKANVAVKV 600
Db 541 LEDKLPHIFGFTLLGDFIYWTDMQRRSIERVHKVKSARDVIIDQLPDLMLGLKANVAVKV 600

Db 541 LEDKLPHIFGFTLLGDFIYWTDMQRRSIERVHKVKSARDVIIDQPLDMLGLKAVNAVW 600
Qy 601 GTNPCADNRGGCSHLCTFTPHATRCGCGIGLELLSDMTKTCIVPBAFLVFTSRAAIIHRL 660
Db 601 GTNPCADNRGGCSHLCTFTPHATRCGCGIGLELLSDMTKTCIVPBAFLVFTSRAAIIHRL 660
Qy 661 ETNNNDVAIPLTGVKEASALDFVSNHHIYWDVSLKTIISRAFMNGSSVEHVFEGLDYP 720
Db 661 ETNNNDVAIPLTGVKEASALDFVSNHHIYWDVSLKTIISRAFMNGSSVEHVFEGLDYP 720
Qy 721 EGMADVWNGKLYWADTGTNRLEIARLQDQFQVILVWRDLNPRSLALDPTKGIYIYTW 780
Db 721 EGMADVWNGKLYWADTGTNRLEIARLQDQFQVILVWRDLNPRSLALDPTKGIYIYTW 780
Qy 781 GKGPRIVRAFMDGTNCMTLVKVGANDLTIDYADQRLYTDLDTNMISSNMLGOERVV 840
Db 781 GKGPRIVRAFMDGTNCMTLVKVGANDLTIDYADQRLYTDLDTNMISSNMLGOERVV 840
Qy 841 IADDLPHFPFGLTQVSDYIYWTDMNLHLSIERADKTSGRNRTLIQHLDFVMDILVFHSSRQ 900
Db 841 IADDLPHFPFGLTQVSDYIYWTDMNLHLSIERADKTSGRNRTLIQHLDFVMDILVFHSSRQ 900
Qy 901 DGLANDCMHNGCCGOLCLAIPEGHRCGCASHYTLDPSSNCSPTTFLPFQKSAISRM 960
Db 901 DGLANDCMHNGCCGOLCLAIPEGHRCGCASHYTLDPSSNCSPTTFLPFQKSAISRM 960
Qy 961 PDDQHSPLIPLHGLRNKVAIDVDPDKFYIYWDGRONIKRAKDDGTOPFVLTSLSOGQ 1020
Db 961 PDDQHSPLIPLHGLRNKVAIDVDPDKFYIYWDGRONIKRAKDDGTOPFVLTSLSOGQ 1020
Qy 1021 NPDRQPHDLSIDIYSRTLFWTCEATNTINVHLSGEAMGVVLRGDRDKPRAIVVNAERG 1080
Db 1021 NPDRQPHDLSIDIYSRTLFWTCEATNTINVHLSGEAMGVVLRGDRDKPRAIVVNAERG 1080
Qy 1081 LYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLFWVDADLRKIESCD 1140
Db 1081 LYFTNMQDRAAKIERAALDGTREVLFTTGLIRPVALVVDNTLGLFWVDADLRKIESCD 1140
Qy 1141 LSGANRLTLEDANIVQPLGLTILGHLYWIDRQOQMIERVKTTGDKRTRIQGRVAHLTG 1200
Db 1141 LSGANRLTLEDANIVQPLGLTILGHLYWIDRQOQMIERVKTTGDKRTRIQGRVAHLTG 1200
Qy 1201 IHAVEVSLEBFAHPCARDNGGCSHICIAKGDGTGPRCSPVHLVLLQNLITCGEPPTCS 1260
Db 1201 IHAVEVSLEBFAHPCARDNGGCSHICIAKGDGTGPRCSPVHLVLLQNLITCGEPPTCS 1260
Qy 1261 PDQFACATGEIDCIPGAWRCDFPECDQSDDEBGCPCVCSAAQFPCARGCQVDLRLRCDGE 1320
Db 1261 PDQFACATGEIDCIPGAWRCDFPECDQSDDEBGCPCVCSAAQFPCARGCQVDLRLRCDGE 1320
Qy 1321 ADCQDRSEADCDALCLPNQFRASGCQVLKQCCDSPDCIDGSDLMCEITKPPSDDS 1380
Db 1321 ADCQDRSEADCDALCLPNQFRASGCQVLKQCCDSPDCIDGSDLMCEITKPPSDDS 1380
Qy 1381 PAHSSAIGFVIGIILSLFVMGGVYFVCQVVCQRYAGANGFPFHEYVSGTGHVPLNFIAP 1440
Db 1381 PAHSSAIGFVIGIILSLFVMGGVYFVCQVVCQRYAGANGFPFHEYVSGTGHVPLNFIAP 1440
Qy 1441 GGSQHGPTGTACGKSMSSVSLMGRRGVPLDRNHVHTGASSSSSSSTKATLYPPIINP 1500
Db 1441 GGSQHGPTGTACGKSMSSVSLMGRRGVPLDRNHVHTGASSSSSSSTKATLYPPIINP 1500
Qy 1501 PPSPATDPSLNMDFYSSNIPATAPRVPYIIRGMAPPPTPCSTDVDCDSYSASRWKAS 1560
Db 1501 PPSPATDPSLNMDFYSSNIPATAPRVPYIIRGMAPPPTPCSTDVDCDSYSASRWKAS 1560
Qy 1561 KYILDNSDSPPPPPPHPSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615
Db 1561 KYILDNSDSPPPPPPHPSQYLSAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615

RESULT 13

US-10-331-907-4

; Sequence 4, Application US/10331907
; Publication No. US20030181660A1
; GENERAL INFORMATION:
; APPLICANT: Todd, John A
; Hess, John W
; Caskey, Charles T
; Cox, Roger D
; Gerhold, David
; Hammond, Holly
; Hey, Patricia
; Kawaguchi, Yoshihiko
; Merriman, Tony R
; Metzker, Michael L
; TITLE OF INVENTION: No. US20030181660A1el LDL-Receptor
; NUMBER OF SEQUENCES: 455
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon and Vanderhye
; STREET: 1100 No. US20030181660A1th Glebe Road, Eighth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: US
; ZIP: VA 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/331.907
; FILING DATE: 31-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/402.923A
; FILING DATE: 14-Feb-2001
; APPLICATION NUMBER: PCT/GB98/01102
; FILING DATE: 15-Apr-1998
; APPLICATION NUMBER: US 60/043.553
; FILING DATE: 15-Apr-1997
; APPLICATION NUMBER: US 60/048.740
; FILING DATE: 05-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: B.J.Sadoff
; REGISTRATION NUMBER: 36,663
; REFERENCE/DOCKET NUMBER: 620-81
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4091
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; TYPE: amino acid
; LENGTH: 1591 amino acids
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-331-907-4

Query Match 98.5%; Score 8607; DB 14; Length 1591;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1590; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 25 CPAPAAASPLLLFANRRDRLVDAGGVKLESTIVVSGLEDAADVDFQFSKGAIVYWDVSE 84
Db 1 CPAPAAASPLLLFANRRDRLVDAGGVKLESTIVVSGLEDAADVDFQFSKGAIVYWDVSE 60
Qy 85 EAIKQTYLNTGTAQVQNVVITSGLYSPDGLACDWDGKLYWTDSETNRIEVANLNGTSKV 144
Db 61 EAIKQTYLNTGTAQVQNVVITSGLYSPDGLACDWDGKLYWTDSETNRIEVANLNGTSKV 120
Qy 145 LFWQDLQDQPAIALDPAHGYWYWDGTPRIERAGMDGSTRKIIIVDSDIYWPNGLTIDL 204
Db 121 LFWQDLQDQPAIALDPAHGYWYWDGTPRIERAGMDGSTRKIIIVDSDIYWPNGLTIDL 180
Qy 205 EEQKLYWADAKLSFIHRANLDGSPRQKVVEGSLTHPPFALTLSGDTLYWTDWQTSIHACN 264
Db 181 EEQKLYWADAKLSFIHRANLDGSPRQKVVEGSLTHPPFALTLSGDTLYWTDWQTSIHACN 240

QY 265 KETGGKKEILSALYSPMDIQVLQSOERQPFHTTRCEEDNGGCSHLCLLSPSPFFTCACP 324
Db 241 KETGGKKEILSALYSPMDIQVLQSOERQPFHTTRCEEDNGGCSHLCLLSPSPFFTCACP 300
QY 325 TGVQLQDNRTCKAGAEVLLARTRDLRRIISLDPDFDIIVLQVDDIRHAIADYDPLE 384
Db 301 TGVQLQDNRTCKAGAEVLLARTRDLRRIISLDPDFDIIVLQVDDIRHAIADYDPLE 360
QY 385 GYVYVTDDEVRAIRAYLDGSAQTLVNTPEINDPGIAVDWVARNLYWTDGTDRIEVTR 444
Db 361 GYVYVTDDEVRAIRAYLDGSAQTLVNTPEINDPGIAVDWVARNLYWTDGTDRIEVTR 420
QY 445 LNGTSRKILVSDLEPRALHPVWGLMYWTDWGENPKIECANLDGQERRVLVNASLGW 504
Db 421 LNGTSRKILVSDLEPRALHPVWGLMYWTDWGENPKIECANLDGQERRVLVNASLGW 480
QY 505 PNLGALDLQEGKLYMGDAKTKIEVINVDGTRKRTLLLEDKLPHPFGFTLLGDFIYWTWQ 564
Db 481 PNLGALDLQEGKLYMGDAKTKIEVINVDGTRKRTLLLEDKLPHPFGFTLLGDFIYWTWQ 540
QY 565 RRSIERVHKVAKASRVDIIDQLPDLMLKAVNVKVVGNPCADNRNGGCSHLCTFFTPHATR 624
Db 541 RRSIERVHKVAKASRVDIIDQLPDLMLKAVNVKVVGNPCADNRNGGCSHLCTFFTPHATR 600
QY 625 CGCPIGLELLSDMKTCIPEAEFLVETSRAAIHRISLETNNNDVAIPLTGVKEASALDFDV 684
Db 601 CGCPIGLELLSDMKTCIPEAEFLVETSRAAIHRISLETNNNDVAIPLTGVKEASALDFDV 660
QY 685 SNHHIYWDVSLKTSRAFNMGSSVEHVVEFGLDYPEGMAVDWGMKNLYWADTGNRIEV 744
Db 661 SNHHIYWDVSLKTSRAFNMGSSVEHVVEFGLDYPEGMAVDWGMKNLYWADTGNRIEV 720
QY 745 ARLDQCFQVLVWRDLNPRSLALDPTKGYIYWTWEGGKPRIVRAFMGTGNTMLVLDKVG 804
Db 721 ARLDQCFQVLVWRDLNPRSLALDPTKGYIYWTWEGGKPRIVRAFMGTGNTMLVLDKVG 780
QY 805 RANDLTIDYADQRLVWTDLTNTMISSNMLGOERVVIAADLPHPEGLTOYSDIYWTWDMN 864
Db 781 RANDLTIDYADQRLVWTDLTNTMISSNMLGOERVVIAADLPHPEGLTOYSDIYWTWDMN 840
QY 865 LHSIERADKTSGRNRTLQGHLDVYMDILVHSSRQDGLNDCMHNNGCGQLCLAIPEGH 924
Db 841 LHSIERADKTSGRNRTLQGHLDVYMDILVHSSRQDGLNDCMHNNGCGQLCLAIPEGH 900
QY 925 RCGCASHYTLDPSSRNCSPPTFTLLFSQSAISRMI PDDQHSPLILPLHGLRNVKAIDY 984
Db 901 RCGCASHYTLDPSSRNCSPPTFTLLFSQSAISRMI PDDQHSPLILPLHGLRNVKAIDY 960
QY 985 DPLDKFIYVWDGRQNIKRAKDDGTQPFVLTSLSQGNPDROPHDLSIDIYSTRTLFWTCEA 1044
Db 961 DPLDKFIYVWDGRQNIKRAKDDGTQPFVLTSLSQGNPDROPHDLSIDIYSTRTLFWTCEA 1020
QY 1045 TINTINVRHLSGEAMGVILRGDRDKPRAI VVNAERGILYFTNMQDRAAKIERAALDGTRE 1104
Db 1021 TINTINVRHLSGEAMGVILRGDRDKPRAI VVNAERGILYFTNMQDRAAKIERAALDGTRE 1080
QY 1105 VLFTTGLIRPVALVVDNTLGLKFWVDADLKRIESCDSLGSANRLTLEDANIVQPLGLTILG 1164
Db 1081 VLFTTGLIRPVALVVDNTLGLKFWVDADLKRIESCDSLGSANRLTLEDANIVQPLGLTILG 1140
QY 1165 KHLVYWDQOQMIERVEKTTGDKRTIOGRVAHLTGTHIAVEVLSLEEPSAHPCARDNGGC 1224
Db 1141 KHLVYWDQOQMIERVEKTTGDKRTIOGRVAHLTGTHIAVEVLSLEEPSAHPCARDNGGC 1200
QY 1225 SHICIAKGDTGTPRCSVHLVLLQNLTCGPPTCSPDQFACATGEIDCIPGAWRCDGFP 1284
Db 1201 SHICIAKGDTGTPRCSVHLVLLQNLTCGPPTCSPDQFACATGEIDCIPGAWRCDGFP 1260
QY 1285 ECDQSDDEEGCPVCSAAGFCARGOCVRLRCDEADQCRSDREADCDACILNPQRCA 1344
Db 1261 ECDQSDDEEGCPVCSAAGFCARGOCVRLRCDEADQCRSDREADCDACILNPQRCA 1320

QY 1345 SGQCVLIKQCCDSPDCIDGSDCEITKTPSDSDSPAHSASAIQPVIGIILSLFWMGVY 1404
Db 1321 SGQCVLIKQCCDSPDCIDGSDCEITKTPSDSDSPAHSASAIQPVIGIILSLFWMGVY 1380
QY 1405 FVQCRVWCORVAGANGPPHYVSGTHVPLNFIAPGSQHGPFTGIAAGKSMSSVSLM 1464
Db 1381 FVQCRVWCORVAGANGPPHYVSGTHVPLNFIAPGSQHGPFTGIAAGKSMSSVSLM 1440
QY 1465 GGRGVPLVDRNHNVTGASSSSSTKATLYPPIILNPPSPATDPSLYNMDMFYSSNIPAT 1524
Db 1441 GGRGVPLVDRNHNVTGASSSSSTKATLYPPIILNPPSPATDPSLYNMDMFYSSNIPAT 1500
QY 1525 ARVPRPYIIRGMAPPTTPCSTDVCDSDYASRWKASKYKYLNDSDSDYPPPPPHSQYL 1584
Db 1501 VRPRPYIIRGMAPPTTPCSTDVCDSDYASRWKASKYKYLNDSDSDYPPPPPHSQYL 1560
QY 1585 SAEDSCPPSPATERSYFHLFPFPPSPCTDSS 1615
Db 1561 SAEDSCPPSPATERSYFHLFPFPPSPCTDSS 1591

RESULT 14

US-10-331-907-43

; Sequence 43, Application US/10331907

; Publication No. US20030181660A1

; GENERAL INFORMATION:

; APPLICANT: Todd, John A

; Hess, John W

; Caskey, Charles T

; Cox, Roger D

; Gerhold, David

; Hammond, Holly

; Hey, Patricia

; Kawaguchi, Yoshihiko

; Merriman, Tony R

; Metzker, Michael L

; TITLE OF INVENTION: No. US20030181660A1e1 LDL-Receptor

; NUMBER OF SEQUENCES: 455

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Nixon and Vanderhye

; STREET: 1100 No. US20030181660A1e1 Glebe Road, Eighth Floor

; CITY: Arlington

; STATE: Virginia

; COUNTRY: US

; ZIP: VA 22201-4714

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/331,907

; FILING DATE: 31-Dec-2002

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: US/09/402,923A

; FILING DATE: 14-Feb-2001

; APPLICATION NUMBER: PCT/GB98/01102

; FILING DATE: 15-APR-1998

; APPLICATION NUMBER: US 60/043,553

; FILING DATE: 15-APR-1997

; APPLICATION NUMBER: US 60/048,740

; FILING DATE: 05-JUN-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: B.J.Sadoff

; REGISTRATION NUMBER: 36,663

; REFERENCE/DOCKET NUMBER: 620-81

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703) 816-4091

; TELEFAX: (703) 816-4100

; INFORMATION FOR SEQ ID NO: 43:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1591 amino acids

; TYPE: amino acid

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 43:

US-10-331-907-43

Query Match 98.5%; Score 8607; DB 14; Length 1591;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1590; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 25 CPAPAAAPLFFANRRDRLVLDAGGVKLESTIVVSGLEDAADVDFQSKGAVYWDVSE 84
DB 1 CPAPAAAPLFFANRRDRLVLDAGGVKLESTIVVSGLEDAADVDFQSKGAVYWDVSE 60

QY 85 EAIKQTYLNQTCGAQVNVVIVSGLSPDGLADWVKLYTWDSETNRIEVANLNGTSKV 144
DB 61 EAIKQTYLNQTCGAQVNVVIVSGLSPDGLADWVKLYTWDSETNRIEVANLNGTSKV 120

QY 145 LFWQDLDPRAIALDPAGHYMYTWDGETPRIERAGMDGSTRKIIIVSDIYWPNGLTIDL 204
DB 121 LFWQDLDPRAIALDPAGHYMYTWDGETPRIERAGMDGSTRKIIIVSDIYWPNGLTIDL 180

QY 205 EEOKLYWADAKLSFTHRANLDSGPRQKVVESGLTHPPALTLTSGDTLYWTDWQTSIHACN 264
DB 181 EEOKLYWADAKLSFTHRANLDSGPRQKVVESGLTHPPALTLTSGDTLYWTDWQTSIHACN 240

QY 265 KRTGGKRKEILSALYSPMDIQVLSQERQPPFFHTRCEEDNGGCSHLCLLSPSEPFYTCACP 324
DB 241 KRTGGKRKEILSALYSPMDIQVLSQERQPPFFHTRCEEDNGGCSHLCLLSPSEPFYTCACP 300

QY 325 TGVQLQNGRTCKAGAEVLLARLTDLRRLSLDTPDFTDVLQVDDIRHAIADYDPLE 384
DB 301 TGVQLQNGRTCKAGAEVLLARLTDLRRLSLDTPDFTDVLQVDDIRHAIADYDPLE 360

QY 385 GYVYVWTDDEVAIRRAYLDGSGAOTLVNTEINDPGIAVWVARNLYWTDGTDRIVETR 444
DB 361 GYVYVWTDDEVAIRRAYLDGSGAOTLVNTEINDPGIAVWVARNLYWTDGTDRIVETR 420

QY 445 LNGTSRKILVSEDLDEPAIRALHPVMGLMYTWDGENPKIECANLDGQERRVLNVLNLSGW 504
DB 421 LNGTSRKILVSEDLDEPAIRALHPVMGLMYTWDGENPKIECANLDGQERRVLNVLNLSGW 480

QY 505 PNGALDLQEGKLYWGDADTKIEVINVDGTYKRTLLBKLPHIFGFTLLGDFIYWDWQ 564
DB 481 PNGALDLQEGKLYWGDADTKIEVINVDGTYKRTLLBKLPHIFGFTLLGDFIYWDWQ 540

QY 565 RRSIERVHKVKAARDVILDLQDLMLGKLVANVAVKVGTPNCDNRNGGCSHLCTFTPHATR 624
DB 541 RRSIERVHKVKAARDVILDLQDLMLGKLVANVAVKVGTPNCDNRNGGCSHLCTFTPHATR 600

QY 625 CGCPIGLELLSDMKTCIVPEAFVFTSRAAIHRSLETNNNDVAIPLTGKVEASALDFDV 684
DB 601 CGCPIGLELLSDMKTCIVPEAFVFTSRAAIHRSLETNNNDVAIPLTGKVEASALDFDV 660

QY 685 SNNHIYWDVSLKTSIRAFMNGSSVEHVVEGLDYPEGMAVDWVGKKNLYWADGTGNRIEV 744
DB 661 SNNHIYWDVSLKTSIRAFMNGSSVEHVVEGLDYPEGMAVDWVGKKNLYWADGTGNRIEV 720

QY 745 ARLDGQFQVILVWDLNDRSLALDPKGYLYWTEWGGKPRIVRAFMDGTNCMTLVKVG 804
DB 721 ARLDGQFQVILVWDLNDRSLALDPKGYLYWTEWGGKPRIVRAFMDGTNCMTLVKVG 780

QY 805 RANDLTIDYADQRLYWTDLDTNMIESNMLQGERVVIADDLPHFPGLTQYSDIYWTWVN 864
DB 781 RANDLTIDYADQRLYWTDLDTNMIESNMLQGERVVIADDLPHFPGLTQYSDIYWTWVN 840

QY 865 LHSIERADKTSGRNRTLQGHLDVMDILVPHSSRQDGLNDCMNNNGCGQLCLAIPEGH 924
DB 841 LHSIERADKTSGRNRTLQGHLDVMDILVPHSSRQDGLNDCMNNNGCGQLCLAIPEGH 900

QY 925 RCGCASHYTLDPSSRNCSPPTTLLFSQKSAISRMIPDDQHSPLILPLHGLRNVKAIDY 984
DB 901 RCGCASHYTLDPSSRNCSPPTTLLFSQKSAISRMIPDDQHSPLILPLHGLRNVKAIDY 960

QY 985 DPLDKFIYVWDRQNIKRAKDDGTQPFVLTSLSQQNPDQPHDLSIDIYSRTLFWTCEA 1044

DB 961 DPLDKFIYVWDRQNIKRAKDDGTQPFVLTSLSQQNPDQPHDLSIDIYSRTLFWTCEA 1020

QY 1045 TINTINVRHLSGEAMGVVLRGDRDPRRAIVNAERGILYFTNMQDRAAKIERAALDGTRE 1104

DB 1021 TINTINVRHLSGEAMGVVLRGDRDPRRAIVNAERGILYFTNMQDRAAKIERAALDGTRE 1080

QY 1105 VLFTTGLIRPVVALVVDNTLGLKFWVDADLKRIESCDLSGANRLTLEDANIYVQPLGLTILG 1164

DB 1081 VLFTTGLIRPVVALVVDNTLGLKFWVDADLKRIESCDLSGANRLTLEDANIYVQPLGLTILG 1140

QY 1165 KHLWYIDRQQQMIBERVEKTTGDKRTRIQRVAHLTGIHAEVEESLEEFSAHPCARDNGC 1224

DB 1141 KHLWYIDRQQQMIBERVEKTTGDKRTRIQRVAHLTGIHAEVEESLEEFSAHPCARDNGC 1200

QY 1225 SHICIAKGDGTGPRSCPVHLVLLQNLTLTCEBPTTSPDQFACATGEIDCIPGAWRCDGFP 1284

DB 1201 SHICIAKGDGTGPRSCPVHLVLLQNLTLTCEBPTTSPDQFACATGEIDCIPGAWRCDGFP 1260

QY 1285 ECDQSBEGGCPVCSAAQFPCCARGQCVDLRLRCDEADQDRSEADCDALCLPNQFRCA 1344

DB 1261 ECDQSBEGGCPVCSAAQFPCCARGQCVDLRLRCDEADQDRSEADCDALCLPNQFRCA 1320

QY 1345 SGQCVLIKQCCDSFPDCIDGSDLMCEITKPPSDSPAHSAGIIPVIGIILSLFVWGGVY 1404

DB 1321 SGQCVLIKQCCDSFPDCIDGSDLMCEITKPPSDSPAHSAGIIPVIGIILSLFVWGGVY 1380

QY 1405 FVQORVVCQRYAGANGFPFHYVSGTTPHVLNFTAPGSGHQGPTGTIACGSKMSSVSLM 1464

DB 1381 FVQORVVCQRYAGANGFPFHYVSGTTPHVLNFTAPGSGHQGPTGTIACGSKMSSVSLM 1440

QY 1465 GGRGGVPLVDRNHVTGASSSSSTKATLYPPIINPPSPATDPSLYNMDMFYSSNIPAT 1524

DB 1441 GGRGGVPLVDRNHVTGASSSSSTKATLYPPIINPPSPATDPSLYNMDMFYSSNIPAT 1500

QY 1525 ARPYRPIIRGMAPPPTTCTSDVCDSDYSASRWKASKYLYDLNDSDDPYPPTPHSOYL 1584

DB 1501 ARPYRPIIRGMAPPPTTCTSDVCDSDYSASRWKASKYLYDLNDSDDPYPPTPHSOYL 1560

QY 1585 SAEDSCPPSPATERSYFHLPPPPSPCTDSS 1615

DB 1561 SAEDSCPPSPATERSYFHLPPPPSPCTDSS 1591

RESULT 15

US-10-331-907-29

; Sequence 29, Application US/10331907

; Publication No. US20030181660A1

; GENERAL INFORMATION:

APPLICANT: Todd, John A

Hess, John W

Caskey, Charles T

Cox, Roger D

Gerhold, David

Hammond, Holly

Hey, Patricia

Kawaguchi, Yoshihiko

Merriman, Tony R

Metzker, Michael L

TITLE OF INVENTION: NO. US20030181660A1el LDL-Receptor

NUMBER OF SEQUENCES: 455

CORRESPONDENCE ADDRESS:

ADDRESSEE: Nixon and Vanderhye

STREET: 1100 No. US20030181660A1th Glebe Road, Eighth Floor

CITY: Arlington

STATE: Virginia

COUNTRY: US

ZIP: VA 22201-4714

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/331,907
FILING DATE: 31-Dec-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/402,923A
FILING DATE: 14-Feb-2001
APPLICATION NUMBER: PCT/GB98/01102
FILING DATE: 15-APR-1998
APPLICATION NUMBER: US 60/043,553
FILING DATE: 15-APR-1997
APPLICATION NUMBER: US 60/048,740
FILING DATE: 05-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: B. J. Sadoff
REGISTRATION NUMBER: 36,563
REFERENCE/DOCKET NUMBER: 620-81
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)816-4091
TELEFAX: (703)816-4100
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 1639 amino acids
TYPE: amino acid
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 29:
US-10-331-907-29

Query Match 98.1%; Score 8574; DB 14; Length 1639;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1585; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY	26	PAPAAASPLLPANRRDRLVDAGVKLESTIVUSGLEDAADVFQFSGAVYVTDVSEE	85
DB	50	PSLHNSPLLPANRRDRLVDAGVKLESTIVUSGLEDAADVFQFSGAVYVTDVSEE	109
QY	86	AIKQTYLNTQGAQVNVISGLVSPDGLACDVGKLYWTDSETRIEVANLNGTSRKVL	145
DB	110	AIKQTYLNTQGAQVNVISGLVSPDGLACDVGKLYWTDSETRIEVANLNGTSRKVL	169
QY	146	FWQDLQDPRALDPAHGMVYWDWGETPRIERAGMDGSTRKIIIVDSIYPNGLTIDLE	205
DB	170	FWQDLQDPRALDPAHGMVYWDWGETPRIERAGMDGSTRKIIIVDSIYPNGLTIDLE	229
QY	206	EQKLYWADAKLSFTHANLDSGFRQKVVESGLTHPEFALTLSGDTLYWTDWQTRSHACNK	265
DB	230	EQKLYWADAKLSFTHANLDSGFRQKVVESGLTHPEFALTLSGDTLYWTDWQTRSHACNK	289
QY	266	RTGGRKEILSALYSPMDIQVLSQERQPFPHTRCEDNGGCSHLCLLSPSFPFYTCAPT	325
DB	290	RTGGRKEILSALYSPMDIQVLSQERQPFPHTRCEDNGGCSHLCLLSPSFPFYTCAPT	349
QY	326	GVQLQDNQRTCKAGAEVLLARLTDLRISLDTPDFDVIQVDDIRHAIADYDPLEG	385
DB	350	GVQLQDNQRTCKAGAEVLLARLTDLRISLDTPDFDVIQVDDIRHAIADYDPLEG	409
QY	386	VYVYTTDDVRAIRRAYLDGSGAQTIVNTEINDPGIAVDWVARNLYWTDGTRIEVTRL	445
DB	410	VYVYTTDDVRAIRRAYLDGSGAQTIVNTEINDPGIAVDWVARNLYWTDGTRIEVTRL	469
QY	446	NGTSRKILVSEDLDEPRALALHPVMGLMYWTDWGNPKIECANLDGQERRVLVNASLQWP	505
DB	470	NGTSRKILVSEDLDEPRALALHPVMGLMYWTDWGNPKIECANLDGQERRVLVNASLQWP	529
QY	506	NGLALDQEGKLYWGDAKTKEVINVDGTRKRTILLEDKLPHIFGTLLGDFIYWDWQR	565
DB	530	NGLALDQEGKLYWGDAKTKEVINVDGTRKRTILLEDKLPHIFGTLLGDFIYWDWQR	589
QY	566	RSIERVHVKAASRDVIIDQLPDLMLKAVNVAKVVGTVNFCADRNGGCSHLFFTPHATRC	625
DB	590	RSIERVHVKAASRDVIIDQLPDLMLKAVNVAKVVGTVNFCADRNGGCSHLFFTPHATRC	649
QY	626	GCPIGLELLSDMKTCIVPEAFIVFTSRAAIIHRIISLETTNNNDVAIPLTGKESALDFDVS	685

DB	650	GCPIGLELLSDMKTCIVPEAFIVFTSRAAIIHRIISLETTNNNDVAIPLTGKESALDFDVS	709
QY	686	NNHIYWTVDVSLKTTISRAPMNGSSVEHVPEGLDYPEGMAVDMGKNLYWATGNNRIEVA	745
DB	710	NNHIYWTVDVSLKTTISRAPMNGSSVEHVPEGLDYPEGMAVDMGKNLYWATGNNRIEVA	769
QY	746	RLDGQFROVLVWRDLNPRSLALDPTKGYIYVWTEWGGKPRIVRAFMDOGTNCMTLVKVR	805
DB	770	RLDGQFROVLVWRDLNPRSLALDPTKGYIYVWTEWGGKPRIVRAFMDOGTNCMTLVKVR	829
QY	806	ANDLTIDYADORLYWTDLTNNIESSNNMLGQERVVIAADLPHFPGLTQYSYIYWTWNWL	865
DB	830	ANDLTIDYADORLYWTDLTNNIESSNNMLGQERVVIAADLPHFPGLTQYSYIYWTWNWL	889
QY	866	HSIERADKTSGRNRTLIGQHLDFVMDILVHFSSRODGLNDCHMNNQCGQLCLAIPEGHR	925
DB	890	HSIERADKTSGRNRTLIGQHLDFVMDILVHFSSRODGLNDCHMNNQCGQLCLAIPEGHR	949
QY	926	CGCASHYTLDPSSRNCSPPTFTLLPSOKSAISRMIPDOHSPDLILPLHGLRNKKAIDYD	985
DB	950	CGCASHYTLDPSSRNCSPPTFTLLPSOKSAISRMIPDOHSPDLILPLHGLRNKKAIDYD	1009
QY	986	PLDKFIYWDGRQNIKRAKODGTQPFVLVLTSLSQCNPDROPHDLSIDIYSRTLFWTCEAT	1045
DB	1010	PLDKFIYWDGRQNIKRAKODGTQPFVLVLTSLSQCNPDROPHDLSIDIYSRTLFWTCEAT	1069
QY	1046	NTINVHRLSGEAMGVLRGRDRKPRAIIVVNAERGILYFTNMQDRAAKIERAALDGTREVE	1105
DB	1070	NTINVHRLSGEAMGVLRGRDRKPRAIIVVNAERGILYFTNMQDRAAKIERAALDGTREVE	1129
QY	1106	LFTTGLIRPALVVDNTLGLKFWVDADLKRITESCDLSANRLTLEDANIVQPLGTLILGK	1165
DB	1130	LFTTGLIRPALVVDNTLGLKFWVDADLKRITESCDLSANRLTLEDANIVQPLGTLILGK	1189
QY	1166	HLVWIDRQQQMIERVEKTDGKTRIQGRVAHLTGIHAAVEEVSLEEFSAHPCARDNGGCS	1225
DB	1190	HLVWIDRQQQMIERVEKTDGKTRIQGRVAHLTGIHAAVEEVSLEEFSAHPCARDNGGCS	1249
QY	1226	HICAKGDTGRPCSPVHLVLLQNLTCGPPPTCSPOFACATGEIDCIPGAWRCDGPPE	1285
DB	1250	HICAKGDTGRPCSPVHLVLLQNLTCGPPPTCSPOFACATGEIDCIPGAWRCDGPPE	1309
QY	1286	CDDQSDDEGCPVCSAAQPPCARGOCVDLRLRCDEADCDQSDRDEADCDALCPNQFRCAS	1345
DB	1310	CDDQSDDEGCPVCSAAQPPCARGOCVDLRLRCDEADCDQSDRDEADCDALCPNQFRCAS	1369
QY	1346	GQCVLIRKQCCDPPDCIDGSDLMCEITKPPSDSPAHSAAIGPVIIGILSLFVMGGVYF	1405
DB	1370	GQCVLIRKQCCDPPDCIDGSDLMCEITKPPSDSPAHSAAIGPVIIGILSLFVMGGVYF	1429
QY	1406	VQWVQVQYAGANGPPPHYVSGTPHVPINFTAPGSGHGPFTGIACGKSMSSVSLMG	1465
DB	1430	VQWVQVQYAGANGPPPHYVSGTPHVPINFTAPGSGHGPFTGIACGKSMSSVSLMG	1489
QY	1466	GRGVPVLYDRNHVTGASSSSSSSTKATLYPILNPPSPATDPSLYNMDFYSSNIPATA	1525
DB	1490	GRGVPVLYDRNHVTGASSSSSSSTKATLYPILNPPSPATDPSLYNMDFYSSNIPATA	1549
QY	1526	RPYRPYIIRGMAPPPTPCSTDVDCSDYSASRWKASKYIYLDNLSDDSPYPPPTPHSOLYS	1585
DB	1550	RPYRPYIIRGMAPPPTPCSTDVDCSDYSASRWKASKYIYLDNLSDDSPYPPPTPHSOLYS	1609
QY	1586	AEDSCPPSPATERSYFHLFPSPSPCTDSS	1615
DB	1610	AEDSCPPSPATERSYFHLFPSPSPCTDSS	1639

Search completed: February 17, 2005, 01:48:28
Job time : 188 secs